LOCATING REGIONS OF INTEREST PRIOR TO X-RAY IMAGING USING STEREO-PHOTOGRAMMETRY

by

Radhika Kulkarni (KLKRAD001)

A thesis submitted to the University of Cape Town in partial fulfilment of the requirements for the degree of MSc (Med) in Biomedical Engineering

Faculty of Health Sciences

University of Cape Town

02 November 2011

Supervisor: Assoc. Prof Tania Douglas

DECLARATION

I, Radhika Kulkarni , hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work or any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed _____

Date _____

ACKNOWLEDGEMENTS

I would like to thank Professor Tania Douglas, my supervisor, for her constant support, patience, and guidance while I was pursuing my Master's degree. I appreciate all her contributions of time, ideas, and funding to make my masters experience productive and stimulating.

Thanks to Dr Tinashe Mutsvangwa for providing some useful insights and suggestions that contributed to the outcome of my project.

I would also like to thank Mr Charles Harris for designing the camera mounting unit for the project, and Dr Fred Nicolls for providing cameras for the project. A special thanks to Mr Stef Steiner of Lodox Systems for providing all the useful information for the project.

I am indebted to many of my colleagues from the Department of Biomedical Engineering who voluntarily allowed me to take their digital images for various experiments conducted in the research project.

I would also like to thank my in-laws, my mother, and my brother for all their love and encouragement. A special thanks to my husband Dr Amod Kulkarni for his support, faith, encouragement, and patience throughout my studies. Thanks to all my friends for their constant and enduring support.

Lastly, I would like to thank Lodox Systems and the national research foundation (NRF; THRIP) for funding the project, and the members of the staff and the students in the Biomedical Engineering division for their support.

ABSTRACT

The research project aims at locating regions of interest (ROIs) on human subjects prior to X-ray imaging on the Lodox Statscan whole body imaging system. This will save imaging time and minimise exposure to radiation. In addition, X-ray images of smaller size will be stored in the database.

Stereo-photogrammetry was used to find the 3D coordinates of landmark points defining the ROI on digital images taken by cameras mounted on the imaging system. The 3D coordinates obtained were used to find the length, breadth and thickness of the ROI. The length and breadth are used to define the ROI to be scanned. The thickness may be used in determining the scanning parameters. The Integrating Vision Toolkit (IVT) was used to develop the application. Corresponding landmark points need to be selected on both the digital images to determine the 3D coordinates of the landmarks. The selection of corresponding landmarks is facilitated by area based matching techniques, namely zero mean normalized cross correlation (ZNCC) and zero mean sum of absolute differences (ZSAD), and by using retro-reflective circular markers on patients' clothing.

Experiments were conducted to test the accuracy of 3D coordinates obtained using stereo-photogrammetry by calculating inter-landmark distances. These distances were compared with their actual distances measured manually. The application was tested on test objects and on a human subject. Results showed that the mean difference between actual and observed distances for test objects and for a human subject was less than 10 mm which is within the permissible limit of 13 mm as specified by the manufacturers of the Statscan system.

The correspondence matching techniques (ZNCC, ZSAD, and circular marker detection) were tested on digital images of five different subjects. 3D coordinates obtained from the landmark points found using these matching techniques were compared with the 3D coordinates calculated by manually selecting the landmark

points. The errors in X, Y, and Z direction were within the permissible limits of 13 mm. The pixel coordinates of the landmark points obtained using the aforementioned matching techniques were also compared with the pixel coordinates of the landmark points selected manually. The mean pixel errors in X and Y directions were (0.5, 0.6) for ZNCC, (0.6, 0.6) for ZSAD, and (1.1, 1.5) for circular marker detection.

It can be concluded that the stereo-photogrammetry based application can successfully be employed to find the length and breadth parameters required for scanning ROIs, and for measuring thickness for optimising the radiation dose.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Study background

In the late 1980's, there was an increase in diamond theft in the diamond trading company De Beers. To solve this problem, the engineers in the firm designed a very low dose X-ray machine in compliance with international radiation exposure guidelines to scan their employees for hidden diamonds regularly (Vaughan, 2008; Pitcher *et al*., 2008).

The same machine also showed great potential for medical diagnosis. The first clinical scanner (Lodox) was installed in the trauma unit of the Groote Schuur hospital in mid-1990s. The X-ray images acquired from the Lodox machine were diagnostically equivalent to those produced by the conventional analogue X-ray machine and the radiation dose was very low compared to that of a conventional analogue X-ray imaging system (Beningfield *et al*., 1999; Vaughan, 2008).

Further research was done to improve the image quality and reduce the overall time for scanning. A new company, Lodox Systems, took over the development of this scanner. In 2002, the Lodox Statscan system was launched. A newer version of the machine has been installed in the trauma/ paediatric/ forensic units in the hospitals in Cape Town and other locations (Vaughan, 2008).

The linear slot scanning X-ray machine developed by Lodox Systems is shown in Figure 1.1. In conventional X-ray systems, the region of interest is exposed at once to

a cone beam of radiation; in Statscan, a thin X-ray fan beam traverses the region of interest while the image is constructed. Reduced scatter as a result of slot scanning in the Statscan system results in lower radiation dosage when compared with conventional radiography (Irving, 2008).

The Statscan can scan the whole body in less than 13 seconds. The machine consists of an X-ray source on a C-arm, an X-ray detector, patient trolley, and a local positioning tool for selecting the scanning region, as shown in the Figure 1.1 (Beningfield *et al*., 2003).

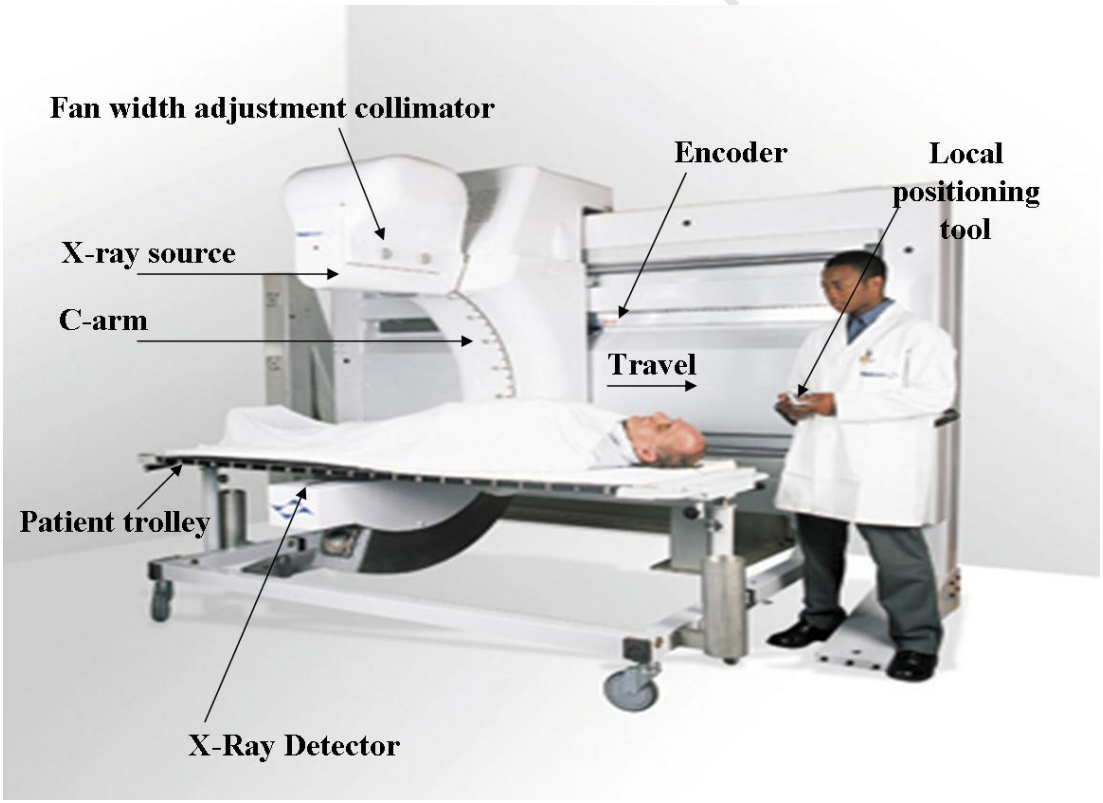

Figure 1.1: Lodox Statscan machine.

The X-ray source produces a collimated fan beam of X-rays as the C-arm moves along the scanning direction (Figure 1.1 and Figure 1.2). The fan beam of X-rays passes through the patient and enters the X-ray detector. The X-rays are detected at the detector end of the C-arm by a bank of scintillator arrays, which convert the X-rays into visible light. These scintillating arrays are optically coupled to charge-coupled devices (CCD cameras) that convert the visible light into digital information resulting in an X-ray image of an object of interest (Beets, 2007).

The X-ray images are constructed while the C-arm is moved linearly along the object of interest (Y-axis in Figure 1.2 and Figure 1.3) and the fan beam of X-rays is perpendicular to the direction of travel (X-axis in Figure 1.2 and Figure 1.3) (Beets, 2007). The rotation of the C-arm around the patient makes it possible to produce X-ray images at different angles (0 to 100 degrees) making antero-posterial, lateral, and oblique imaging possible (Irving *et al.*, 2008; Maree *et al.*, 2007).

The Lodox Statscan machine is capable of scanning the whole body as well as smaller regions of interest (ROI). The patient's details such as size (infant, medium, small, large, extra-large) are entered using the operator console as shown in Figure 1.4. Pre-set scan parameters are selected for the chosen size of the patient. There are seven pre-set trauma procedures and seven general defined procedures. The operator selects the desired procedure. The settings of the X-ray generator (kV and mA) are automatically adjusted based on the selected size and the procedure.

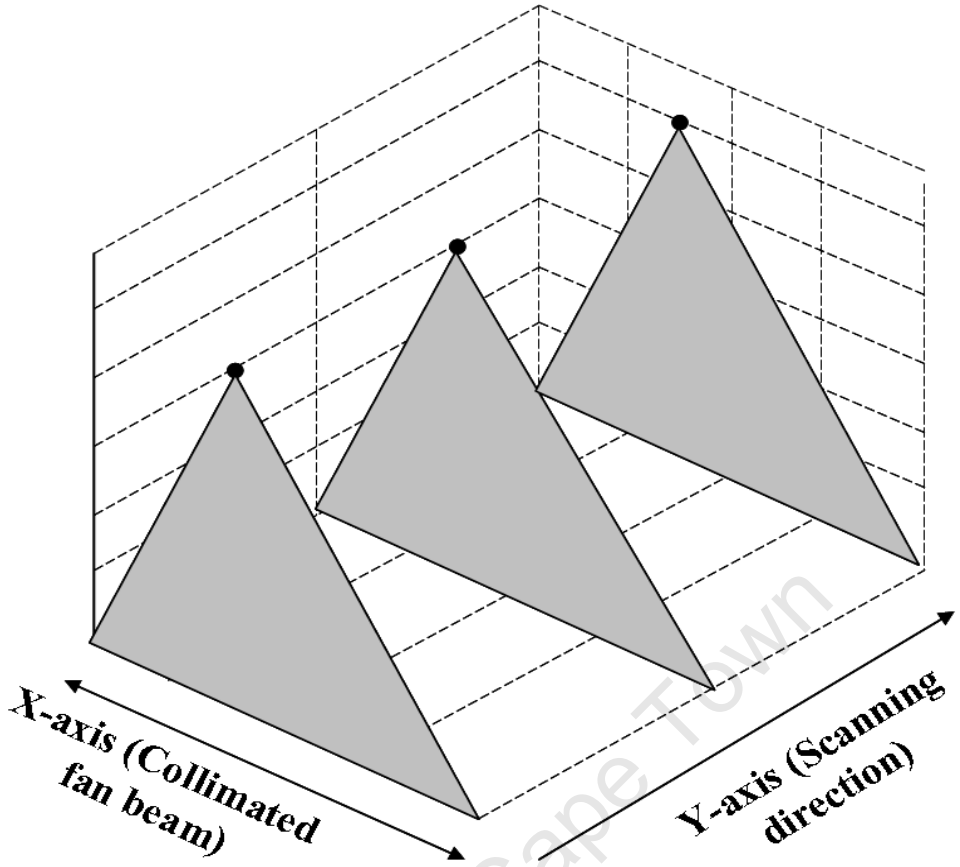

Figure 1.2: Collimated fan beam of X-rays from X-ray source. The X-axis represents the direction of the collimated fan beam and the Y-axis represents the scanning direction or direction of travel of the C-arm (adapted from (Beets, 2007)).

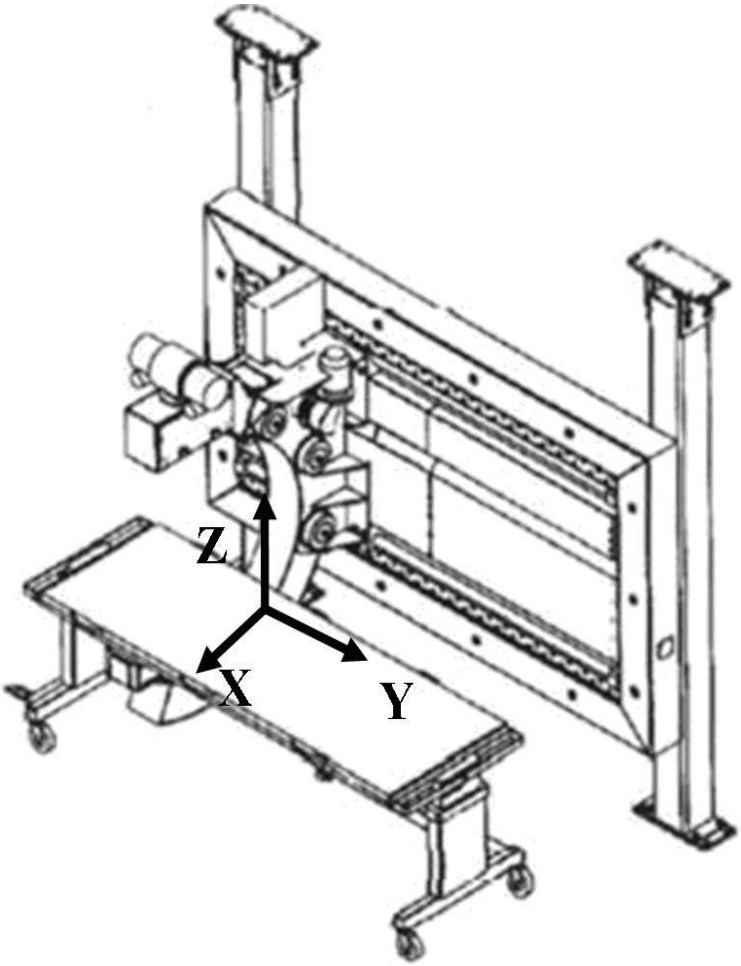

Figure 1.3: Statscan machine with principal axes of operation. The X-axis represents collimator direction, the Y-axis represents the scan direction, and the Z-axis represents direction of the X-ray beams (reproduced from (Beets, 2007)).

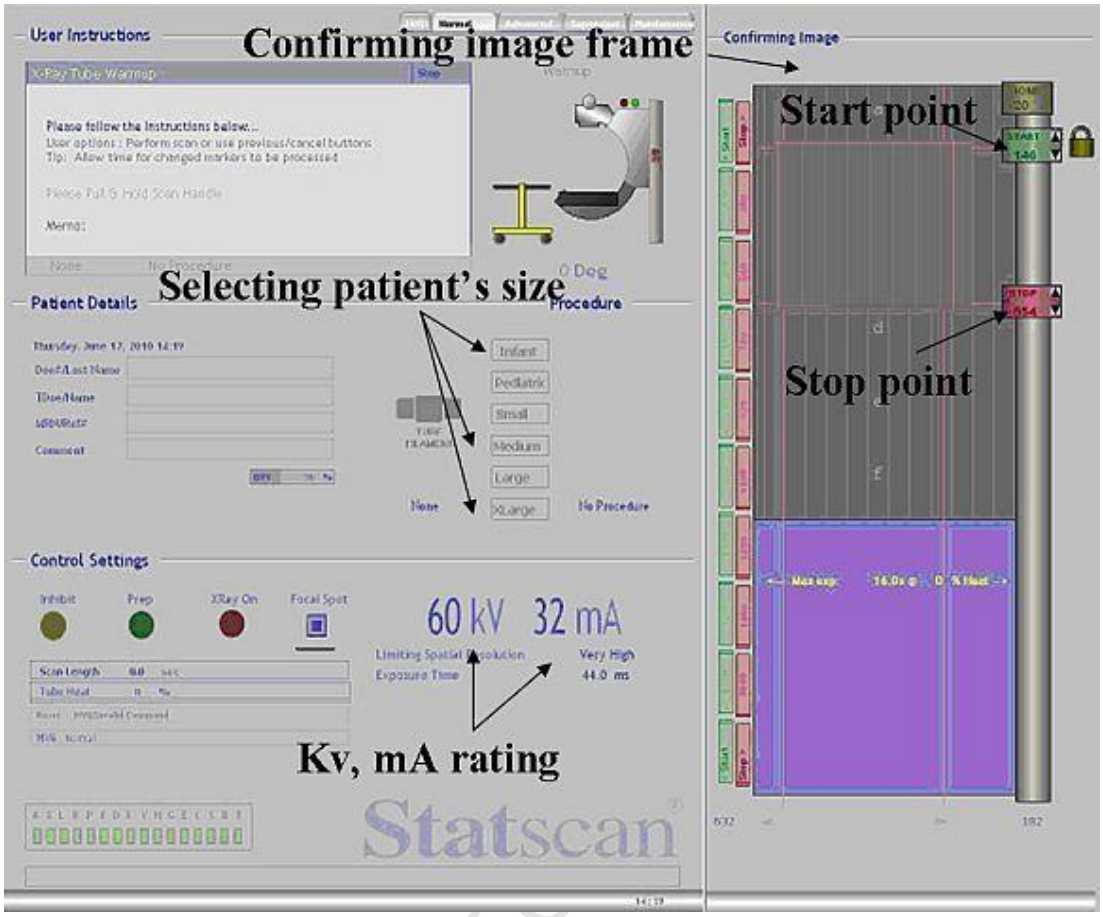

Figure 1.4: Operator console with ROI selected manually using the local positioning tool. The start and stop points are set manually using the local positioning tool by visualizing a laser beam along the direction of travel of C-arm. The patient's image (position) is not available on the confirming image frame prior to the scanning process.

The ROI is currently selected by viewing a laser fan beam along the direction of travel of C-arm, while the C-arm is moving, prior to X-ray imaging. The operator defines the start and stop point for scanning the ROI on the operator console (Figure 1.4) using the local positioning tool shown in Figure 1.1. The start point is marked by the user as the C-arm and laser beam traverse the patient. Once the laser beam has passed through the ROI, the operator marks the stops point manually by clicking the stop button on the local positioning tool. The start and the stop point are stored on the confirming image frame (Figure 1.4). The C-arm is moved to the ROI from its initial

position and the scanning process starts. The scanning process stops automatically once the C-arm crosses the stop point marked earlier by visualizing the laser beam. Once the scanning process is over, the X-ray image of a smaller ROI appears on the confirming image frame shown in Figure 1.5.

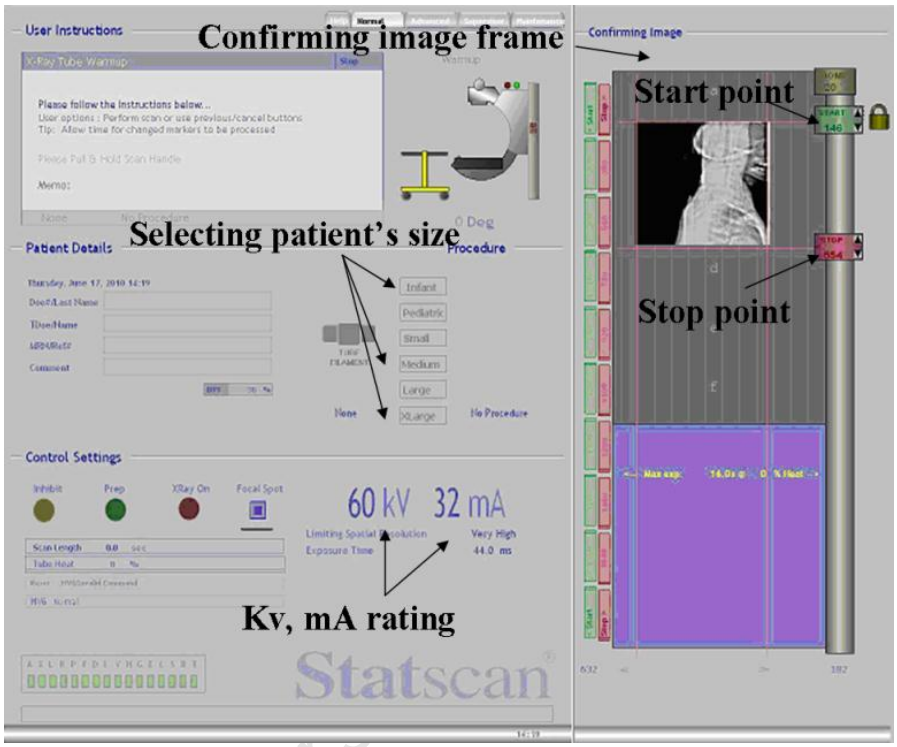

Figure 1.5: Operator console for the Lodox Statscan machine after the ROI is scanned by the C-arm. The X-ray image of the ROI appears on the confirming image frame once the scanning process is over.

The start and the stop points can be set directly on the confirming image frame if the patient's position is known, if a whole body image has already been taken and appears on the confirming image frame (Figure 1.5). Obtaining a whole body image or smaller image for the purpose of determining an ROI would expose the patient to unnecessary radiation.

1.2 Problem statement

Versascan is a vertical, low dose, full body digital X-ray machine being developed by Lodox Systems in order to make the existing machine more versatile. The design of the machine is similar to that of the Scannex machine (shown in Figure 1.6) originally developed by De Beers to detect diamond theft. Figure 1.6A shows the closed box unit and Figure 1.6B shows that the X-rays from the X-ray source projected on the object between the source and the detector. The Versascan will operate similarly to the Scannex machine, in that the patient stands inside the closed box unit and the scanner moves in the vertical direction resulting in a whole body X-ray image of the patient.

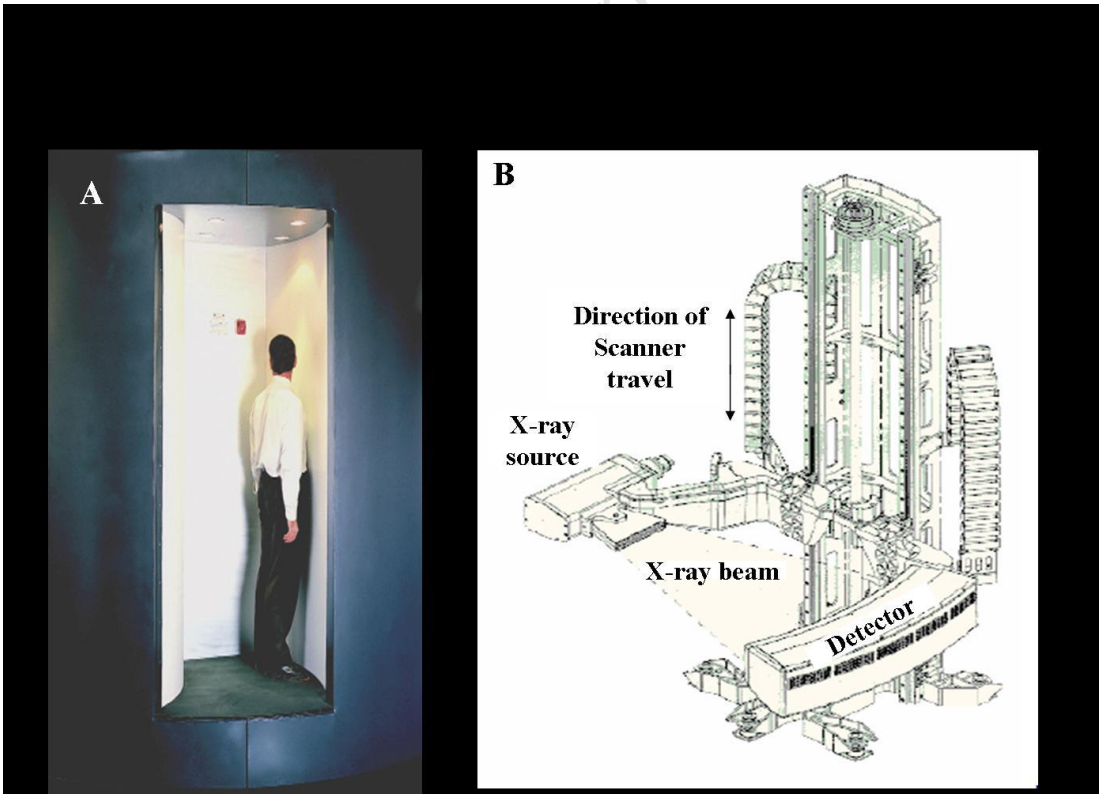

Figure 1.6: Scannex machine. 1.6A shows the closed box unit while B shows that the X-ray energy generated from the X-ray source is projected on the detector while the C-arm moves in the vertical direction (reproduced from (Greenway, 2004)).

However, whole body X-ray imaging is not always desirable. Sometimes, the radiologist may wish to view and store the X-ray image of a smaller ROI. Whole body scanning in such cases will result in exposing patients to unnecessary radiation. In addition, the process of manually splitting the whole body image to view smaller regions of interest is time consuming. The size of the image database will also increase, as larger images are stored in database. This will decrease the image retrieval/viewing speed and efficiency. In order to solve these problems, the functionality of scanning smaller ROIs should be incorporated in the Versascan machine.

The design of the Versascan is similar to that of the Lodox Statscan machine, although Versascan is an upright scanner while Statscan is a horizontal scanner. As mentioned in Section 1.1, in the Statscan machine, the smaller ROI is selected by viewing the laser beam along the direction of travel of C-arm. Since Versascan is a closed box unit, the operator will not be able to view the laser beam on the patient during scanning (in order to define the start and stop point on the operator console). In addition, the laser is expensive to purchase and maintain. Recalibration of the laser beam for accurate imaging is difficult and it requires operator skill to adjust the laser beam to pass exactly through the narrow slit opening for X-ray radiation. Therefore, a better way of finding the ROI for an upright scanner needs to be developed.

Landmark points obtained from digital images of the object of interest can be used to define ROIs as shown in Figure 1.7. The rectangular region in Figure 1.7A shows the ROI and the crosses on Figure 1.7B are the landmark points defining the ROI. The

three-dimensional (3D) coordinates of these landmark points can be estimated and further used to determine length and breadth of the ROI for scanning and thickness for the determination of scanning parameters.

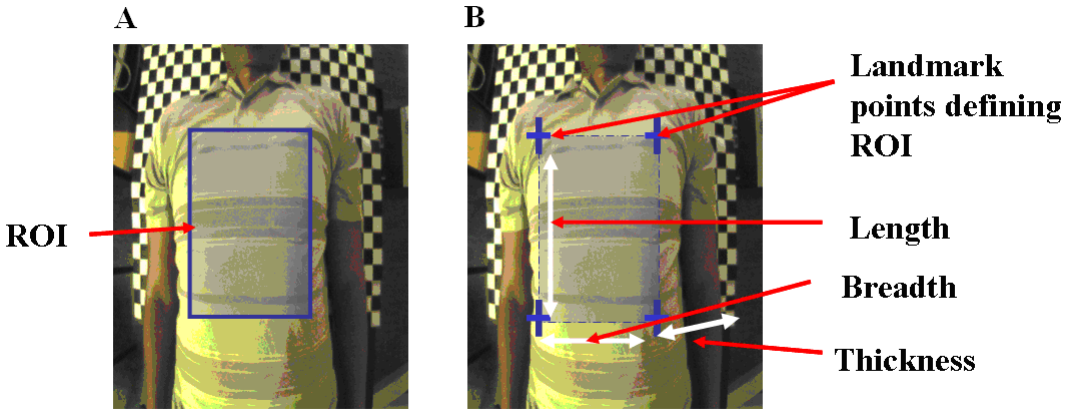

Figure 1.7: ROI defined by landmark points on the object of interest. (A) Patient with the ROI indicated by a rectangular region. (B) Landmark points (shown as cross) can be used to define the ROI.

According to the Food and Drug administration (FDA) (FDA, 1982), the X-ray field dimensions (area exposed to radiation) should correspond to the field size indicated by the beam limiting device (the collimator here). The maximum difference/error in the areas indicated by the X-ray field and the beam limiting device should be within +/- 2 % of the source to image distance (SID) in both the scan and collimator directions (Y and X in Figure 1.2 and Figure 1.3) when the beam axis (Z axis in Figure 1.3) is perpendicular to the plane of the image receptor. The image receptor is the detector with scintillating arrays to convert X-rays into visible light in this case. There might be two sources of error while imaging the ROI. The ROI might not be recognized correctly or the control system in the machine may not correctly

illuminate the indicated ROI with X-ray radiation. If half the errors are allocated to the measurement and half to the control system, the maximum error permitted by the FDA in defining the distances in any linear direction should be within 1% of SID. The SID (for Versascan and Lodox Statscan) is approximately 1300 mm. Thus, the maximum error permitted in calculating the linear distances (length and breadth) of the ROI should be less than or equal to 13 mm (1% of SID). Knowledge of the thickness of the ROI may be used to determine X-ray scanning parameters. While the FDA specifies no permissible thickness measurement error, it is assumed that the error permitted in calculating the thickness of the ROI should also be less than or equal to 13 mm.

1.3 Objectives

The objectives of the research project are:

- To develop a method to locate regions of interest for X-ray imaging on the Lodox Statscan machine based on digital photographs, as illustrated in Figure 1.8.
- To incorporate the method into a computerized tool to accompany the Statscan system.
- To test the application on the Statscan machine.

A tool that is successful on the Statscan machine may later be incorporated into Versascan; however, this was beyond the scope of the project.

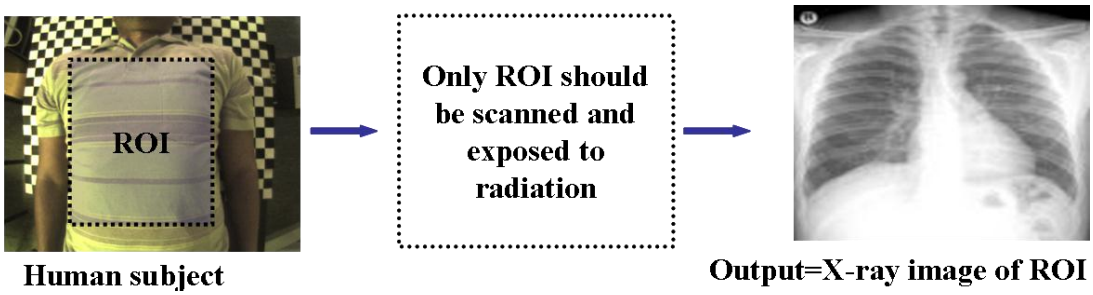

Figure 1.8: The ROI should be located prior to X-ray imaging using landmark points obtained from digital images of the object of interest.

1.4 Project summary

The project presents a cost effective, easy to use computerised tool for determining the ROI defining the physical volume to be exposed to radiation prior to the X-ray imaging process on the Lodox Statscan machine. The landmark points obtained from digital images of the object of interest define the ROI. Stereo-photogrammetry is used to determine the 3D coordinates of the landmark points. The 3D landmark points are then used to calculate various distances defining the physical volume occupied by human subjects.

The computerized tool was developed in the C++ environment using the Integrating Vision Toolkit (IVT) and the OpenCV image processing libraries.

1.5 Dissertation structure

Chapter 2 gives an overview of different techniques for locating ROIs by either utilizing X-ray images or digital images of the object taken by cameras.

Chapter 3 describes the stereo-photogrammetry technique used to find the 3D coordinates of the landmark points using two or more digital images. The steps involved in stereo-photogrammetry such as camera calibration, errors due to distortion, and triangulation are discussed.

Chapter 4 is dedicated to a literature review on different types of camera calibration techniques.

Chapter 5 presents a literature review on different image matching techniques, and their importance in selecting corresponding landmark points in stereo-photogrammetry. A literature review specific to area-based matching techniques and feature based circular marker detection is also presented. Constraints such as epipolar geometry, disparity range, and rectification are covered.

Chapter 6 deals with a preliminary experiment conducted to test stereo-photogrammetry for the proposed solution using the available cameras, an existing 3D calibration frame, and Australis software.

In **Chapter 7**, the research methodology adopted for the proposed solution is presented. The materials, software, and details of the method developed to find the length, breadth, and the thickness of ROI using stereo-photogrammetry are described. The image matching techniques to improve landmark selection are also given.

Chapter 8 describes the solution developed using stereo-photogrammetry.

Chapter 9 presents the details of the experiments conducted and results obtained in finding the accuracy of different methods described in Chapter 7.

Chapter 10 presents an overall discussion of the experiments conducted and the results obtained in Chapter 9.

Chapter 11 draws conclusions and makes recommendations.

Some of the work presented appears in the following conference paper:

Kulkarni, R., Mutsvangwa, T., Douglas, T. 2010, "Locating regions of interest prior to X-ray imaging using stereo-photogrammetry". *Proceedings of the 21st Annual Symposium of the Pattern Recognition Association of South Africa*, Stellenbosch, South Africa, http://www.prasa.org/proceedings/2010/prasa2010-23.pdf (Accessed on 10 Feb 2011).

2 OVERVIEW OF METHODS FOR ROI LOCATION

In this chapter, existing methods of locating ROIs are discussed. Some of the methods and applications are for extracting ROIs from X-ray images while some find ROIs on digital still and video images of the human body. Not all the applications and methods discussed in this section are related to locating ROIs prior to X-ray imaging, but might prove helpful in designing a solution for X-ray imaging.

2.1 Extracting ROIs from X-ray images

Different methods are available for segmenting X-ray images into smaller ROIs. Most of these methods involve human intervention and work efficiently for segmenting specific ROIs such as lungs (Li *et al*., 2001), clavicles (Van Ginneken *et al*., 2006), ribs (Plourde *et al*., 2006) and tuberculosis (TB) lesions (Koeslag, 2002) mostly on chest radiographs. These methods may not work efficiently for regions other than those specified.

Rigid template matching involves searching the entire X-ray image for the best match between the image data and a template of the feature defining the ROI. This technique was used by Koeslag (2002) for finding TB lesions in lung radiographs. In some methods, analysis of the edges of the ROI and the intensity differences at the edges are used to locate the ROI (Li *et al*., 2001). Using active contour models or snakes, the segmentation of the X-ray image into smaller ROIs is performed by fitting a curve to the boundary of the ROIs in the X-ray image. An energy function is

used to measure the fitness or accuracy of the curve (Kass *et al*., 1988; Yue *et al*., 1995). Active shape models are similar to active contour models but the search is based on the shape of the ROI (Cootes *et al*., 1995). Active appearance models require a manually annotated set of training sets of the area to be search and the expected ROI is searched for by comparing the shape and texture of the X-ray image (Edwards *et al*., 1998).

The above methods rely on the acquisition of a larger X-ray image from which ROIs are then extracted. They would not be suitable for the current project, as they would suffer from the disadvantages of subjecting patients to additional radiation and requiring more image acquisition time.

2.2 Determining ROIs using two or more digital images

Photogrammetry generally refers to a technique for taking measurements from photographs. If multiple camera orientations and images are involved, 3D landmark coordinates may be obtained, and the technique is called stereo-photogrammetry (Luhmann *et al*., 2006). Stereo-photogrammetry has been used to detect fetal alcohol syndrome (FAS) in children based on 3D landmark-based facial shape analysis (Douglas and Mutsvangwa, 2010). Savara (1965) proposed stereo-photogrammetry for a quantitative, 3D description of tooth and face morphology. Baroni et al (2000) proposed the use of opto-electronics and close range photogrammetry for automatically checking the position of patients in radiotherapy units. The system developed by Baroni et al (2000) computed 3D coordinates of the multiple markers

placed on patient skin (at various anatomical landmarks). Patient alignment and position monitoring was carried out by comparing the 3D positions of the markers (obtained by close-range photogrammetry) with those of the initial reference positions of the markers acquired during the stimulation procedure and/or the first irradiation session. Eriksson et al (1979) used stereo-photogrammetry for measurement of leg ulcers. Stereo-photogrammetry was used to determine parameters such as edge length, surface area, and volume of the leg ulcers. The clinical evaluation of the healing of leg ulcers were then compared with the stereo-photogrammetric parameters. Ferrario et al (1997) studied the growth and development of the nose by measuring 3D coordinates of various nasal landmark points using stereo-photogrammetry. These 3D coordinates of the nasal landmark points were later used for calculating various nasal lengths, which indicated nasal development, nasal volume, and external surface. Takamoto and Schwartz (1986) proposed to use photogrammetry for calculating the thickness of the retinal nerve fibre layer (NFL) over the major vessels by finding 3D coordinates of the landmark points indicating NFL thickness. This technique could be used as an early indicator of pending visual field loss in glaucoma and for monitoring effects of medication and therapy for ocular hypertension. Tredwell et al (1995) developed a method for analysing scoliosis using stereo-photogrammetry. 3D coordinates of the landmarks on each vertebra at each stage of the operation (spinal instrumentation procedure) were computed using stereo-photogrammetry. The 3D changes in the position of the landmark points were studied during the first and last stage of the operation. In addition, vertebral translations and rotations were also computed from the 3D coordinates of the landmark points for analysing scoliosis during the operation.

In stereo-photogrammetry, the measurements are taken without contacting the object or area of interest and therefore the risk of deforming the area of interest can be avoided. The technique can be used for objects of all sizes (Fryer *et al*., 2007) and can be implemented using simple apparatus (i.e. cameras), with fast data collection (Fryer *et al*., 2007; Mitchell and Newton, 2002).

3 STEREO-PHOTOGRAMMETRY

The aim of the project is to locate ROIs prior to X-ray imaging. Landmark points obtained from photographs can be used to define ROIs on objects (Figure 1.7). Stereo-photogrammetry is a process of determining 3D coordinates of points in space using two or more digital images (Fryer *et al*., 2007; Mitchell and Newton, 2002).

In this chapter, the fundaments of a camera system with reference to different coordinate systems are given. The different steps required in obtaining 3D coordinates of the landmark points from the digital images of the ROI are also discussed.

3.1 Fundamentals of the camera model

In this section, some of the terms used in stereo-photogrammetry are introduced and coordinate systems are defined. As mentioned before, stereo-photogrammetry is the process of converting data from two-dimensional (2D) digital images to 3D data. When the 2D digital images are taken from the cameras, there is a loss of information due to a shift from 3D object space to 2D digital images. This loss of information can be caused by the loss of visibility of a particular region from the cameras and geometric changes due to the relative position of cameras with respect to the object, lens distortions, and colour differences between the object in 3D space and in the 2D digital images (Luhmann *et al*, 2006; Mutsvangwa, 2009). Therefore, the geometric and environmental conditions that existed at the time of imaging play

an important role when retrieving 3D information back from the 2D digital images in photogrammetry.

Perspective projection is the geometric basis of photogrammetry. It is defined as the transformation of data from 3D to 2D (Mikhail *et al*., 2001; Mutsvangwa, 2009). The perspective projection can be defined using a pinhole camera. The central perspective pinhole camera model in the positive position is shown in Figure 3.1. The key features are defined as follows (Azad *et al*., 2008):

- The principal axis is a straight line passing through the projection centre Z and perpendicular to the image plane shown in Figure 3.1.
- The principal point is a point where the principal axis meets the image plane.
- The image coordinate system is a 2D coordinate system with origin at the top left corner of the image; its units are pixels.
- The camera coordinate system is a 3D coordinate system with origin at the projection centre Z. The units are in millimetres.
- The world coordinate system is a 3D coordinate system lying anywhere in space. The units are in millimetres.

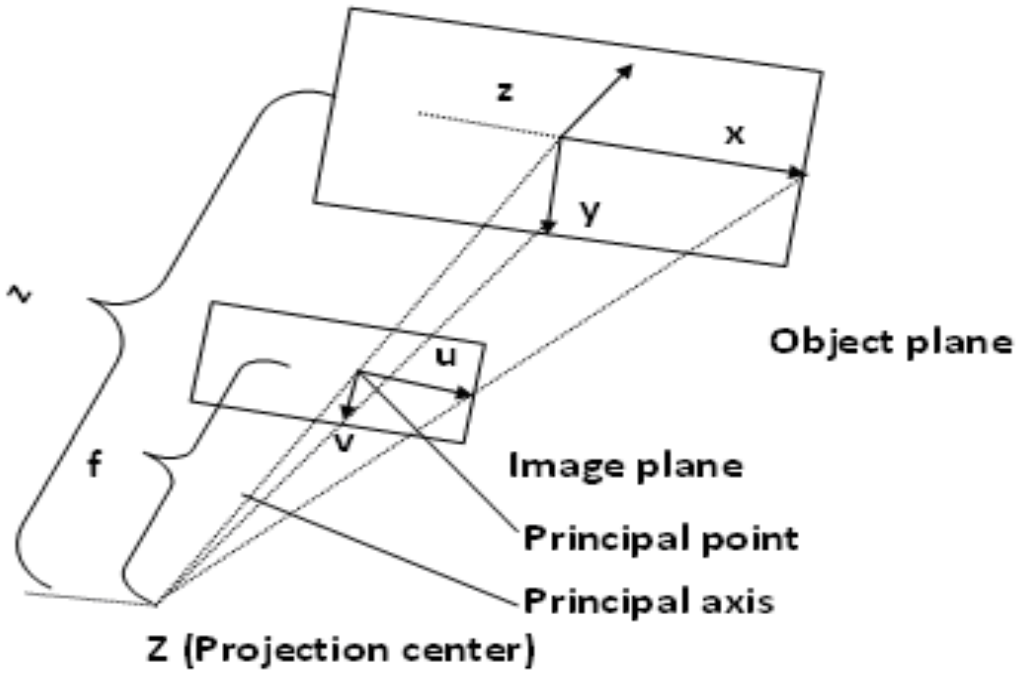

Figure 3.1: The central perspective in a pinhole camera model in the positive direction (adapted from (Azad *et al.*, 2008)).

3.2 Image acquisition

The choice of the camera depends on the object whose 3D coordinates are to be determined. The cameras should have adequate resolution for the task and small radial and tangential distortion (described in Section 3.4). The lighting conditions in the room should be even so that the ROI is clearly visible in both digital images.

3.3 Intrinsic and extrinsic camera parameters

The first step in finding the 3D coordinates of landmark points is to calibrate cameras. Camera calibration is the process of determining the intrinsic and extrinsic parameters of the cameras (Azad *et al.*, 2008).

The intrinsic parameters of the camera are the principal distance and the image coordinates of the principal point (shown in Figure 3.1) (Luhmann *et al*., 2006; Azad *et al*., 2008). These parameters are independent of the choice of world coordinate system i.e. these parameters are fixed unless the focus and the focal length of the camera are changed. In addition, errors due to distortion are also determined. The extrinsic camera parameters transform the world coordinate system to the camera coordinate system (Azad *et al*., 2008).

The extrinsic parameters of the camera are related to rotation and translation as shown in Figure 3.2. Translation is related to the change of position of origin of the 3D coordinate system from the initial position to the new position (Bradski and Kaehler, 2008). Rotation is related to rotation of the 3D coordinate system in each X, Y, and Z plane. The extrinsic parameters must be calculated each time there is a change in the position of cameras as they determine the camera's orientation and position in space (Azad *et al*., 2008). Bradski and Kaehler (2008) determined the intrinsic and extrinsic parameters using a method by Zhang (2000). The mathematical details of the intrinsic and extrinsic parameters are not in the scope of the project and may be found in (Bradski and Kaehler 2008). The different techniques of camera calibration are given in Chapter 4.

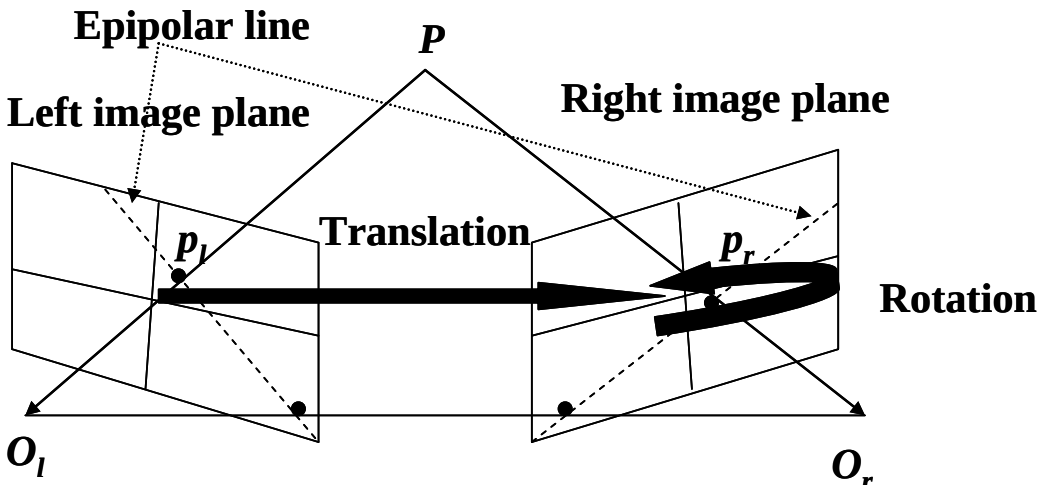

Figure 3.2: Extrinsic parameters of camera (rotation and translation) which describes the relative position of the cameras with respect to each other (adapted from (Bradski and Kaehler, 2008)).

3.4 Errors due to radial and tangential distortion

Errors might occur in the estimation of the intrinsic and extrinsic parameters of the camera due to lens distortion (Azad *et al.*, 2008). Errors might be induced in identifying certain aspects of the object and in finding local features in the image due to lens distortion producing distorted images of the object of interest. If distorted images are used in calculating inter-landmark distances (length, breath, and thickness of the ROI), the distances calculated might not be accurate (Utkarsh, 2011). Therefore, correcting the distortion is important.

Image distortion due to lens distortion is not always uniform. In most cases, the distortion is least at the centre and it increases towards the edges (Montabone, 2010). Radial and tangential distortions are the major sources of error (Bradski and Kaehler, 2008). Therefore, the discussion is restricted to radial and tangential distortion.

Radial distortion occurs due to the shape of the lens and distorts the location of the pixels near the edges of the image as shown in Figure 3.3. Straight lines in the scene appear bent in the image. This effect is clearly visible at the edges of the objects in the image. The image is equally distorted in each direction. The farther the pixels are from the centre, the larger the distortion (Montabone, 2010).

Radial distortion could be barrel or pincushion in nature (Montabone, 2010). In barrel distortion, the corners of the image are bent towards the centre (causing the image to look like a barrel) as shown in Figure 3.3(b). In pincushion distortion, the corners are bent away from the centre as shown in Figure 3.3(c). Bradski and Kaehler (2008) used two parameters for radial distortion, but three for highly distorted cameras.

Tangential distortion occurs due to manufacturing defects. The lenses are not parallel to the imaging plane in this case. Tangential distortion is characterised by two parameters (Bradski and Kaehler, 2008). Zhang (2000) and Brown (1971) have suggested methods to calculate radial and tangential distortion, which need to be taken into account when calculating intrinsic and extrinsic camera parameters. Bradski and Kaehler (2008) used Brown's (1971) method for lens distortion using OpenCV as a digital library.

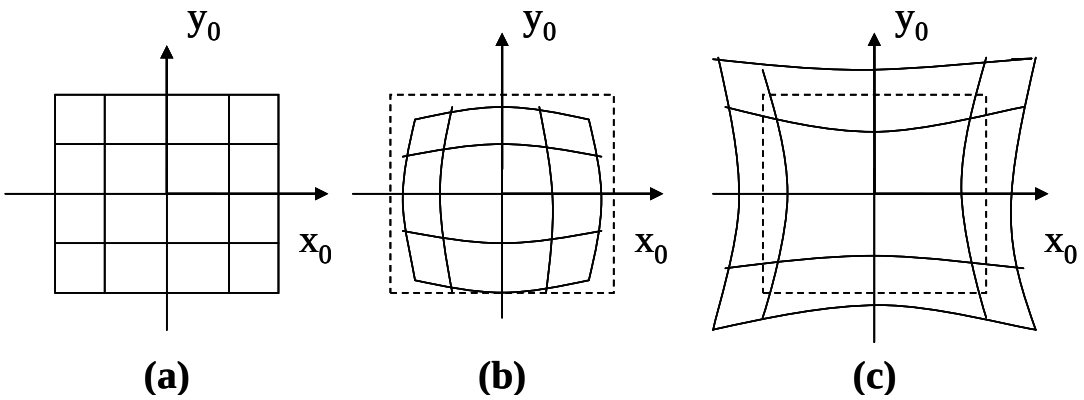

Figure 3.3: Radial distortion: (a) image with no distortion, (b) barrel distortion, (c) pincushion distortion (adapted from (Cao *et al.*, 2010)).

3.5 Triangulation

Once the cameras are calibrated, the next step is to calculate the 3D coordinates of the landmark points using two or more digital images of the object of interest.

The 3D coordinates of landmark points may be calculated using the triangulation technique given in Azad et al (2008). Let the two cameras have the extrinsic parameters 'R_l', 'T_l' (for the left side camera) and 'R_r', 'T_r' (for the right side camera), and the calibration matrices 'C_l' and 'C_r' containing the intrinsic parameters for left and right side cameras respectively. The point 'P' shown in Figure 3.4 whose 3D coordinates have to be determined has image coordinates with pixel values p_l = ($x1$, $y1$) and p_r = ($x2$, $y2$) on the two digital images. All the 3D coordinates mapped to 'p_l' on left digital image will lie on a straight line. Let this line be 'L_l'. Similarly, 'L_r' is a straight line on which all the 3D points mapped to 'p_r' will lie. The camera calibration parameters are used for determining 'L_l' and 'L_r'.

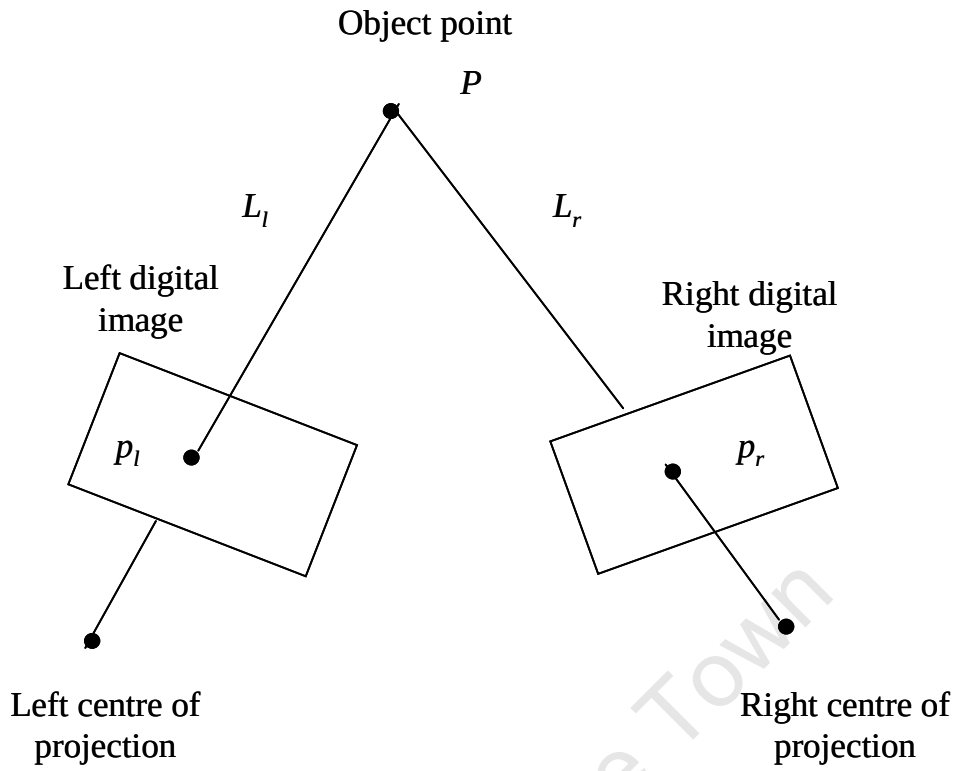

Figure 3.4: Triangulation technique (adapted from (Owens, 1997)).

The intersection of these straight lines ideally represents the 3D coordinates of the point '*P*' in the world coordinate system. However, sometimes, the straight lines might not intersect accurately at one point due to the inaccuracies and the discretization of the pixels of the image sensor (Azad *et al*., 2008). Therefore, the point '*P*' having a minimum distance from both the straight lines ('L_l' and 'L_r') should be searched for. A method of least squares could be used to determine the optimal solution (Azad *et al*., 2008).

3.6 Bundle adjustment

Bundle adjustment is another method of estimating 3D coordinates of landmark points. Bundle adjustment refers to the "bundles" of rays of light leaving all the 3D

points on the object and converging at the centre of each of the cameras. These bundles are adjusted optimally with respect to 3D points (features) on objects and camera position. Bundle adjustment deals with the problem of refining a visual reconstruction to produce optimal 3D structure, camera position, and/or calibration. The optimal estimate for parameters is produced by minimizing some cost function that increases the model fitting error and the solution is optimal for both the 3D structure and viewing parameters (camera position and/or calibration) (Triggs *et al*., 1999). Classically, bundle adjustment computations are looked upon as non-linear least squares (Brown, 1971; Granshaw, 1980). Bundle adjustment is computationally expensive and requires initial approximation of camera orientation parameters for convergence (Shum *et al*., 1999).

3.7 Direct linear transform (DLT)

The direct linear transformation (DLT) can be used for determining 3D coordinates, camera calibration, and for lens distortion correction (Azad *et al*., 2008; Luhmann *et al*., 2006). The method requires a minimum of six reference points from a 3D calibration frame (like the one shown in Figure 4.1) to determine eleven DLT parameters (Azad *et al*., 2008; Luhmann *et al*., 2006). The DLT uses linear equations to determine the orientation of the image. The DLT has an excess of parameters if the parameters of interior orientation are known beforehand. If all the reference points lie on a common plane, it results in a weakly conditioned system of equations. It is not possible to detect the causes of errors in measurement of the image

coordinates and the reference points in the DLT (Azad *et al*., 2008; Luhmann *et al*., 2006).

4 CAMERA CALIBRATION TECHNIQUES

The different techniques for camera calibration using no calibration frame, and one, two and three-dimensional (1D, 2D and 3D) calibration frames, are given below.

4.1 Self-calibration (requires no calibration object)

Self-calibration does not require any calibration frame for camera calibration. The technique uses images of the static object under different conditions for determining camera parameters (intrinsic and extrinsic). The parameters of interior orientation are determined at the time of object measurement. Therefore, if the same camera is used to take all the images, the internal parameters will remain the same for all the images. The external and internal parameters are calculated by measuring the correspondence between a minimum of three images of the object at different orientations. The self-calibration technique can provide an optimal solution for 3D reconstruction of objects but the process involves many parameters and calculations, therefore the results are not always reliable (Luhmann *et al*., 2006; Zhang, 2000).

4.2 Camera calibration using a 3D calibration object

A 3D calibration object has two or more planes (X, Y, Z planes) orthogonal to each other (as shown in Figure 4.1). Circular landmarks are commonly used control points in calibration. The relationship between the 3D object coordinates and the image coordinates is determined by solving for the unknown parameters of the camera model.

Heikkila (2000) proposed a camera calibration model consisting of eight intrinsic parameters and six extrinsic parameters. The radius, 3D coordinates and observed image coordinates of the control points were known in advance. An iterative search method was used to determine the camera parameters. The initial values for camera parameters were solved for by assuming no lens distortion and the rest of the parameters were determined using a linear estimation technique.

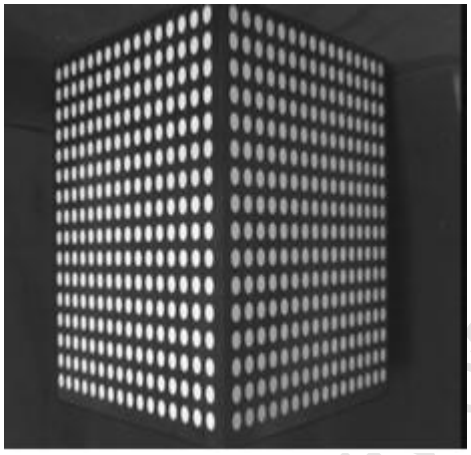

Figure 4.1: Three-dimensional calibration object (reproduced from (Heikkila, 2000)).

4.3 2D plane-based calibration

Zhang (2000) proposed a technique for easily calibrating one or more cameras. The technique is more flexible than the classical techniques of calibrating cameras (such as self-calibration and 3D object calibration). A planar object was used to calibrate the camera. The planar object consisted of a checkerboard black and white pattern as shown in Figure 4.2.

Either the camera or the planar object was moved in different orientations to capture images of the model. The corners of the checkerboard pattern were used as unique features and the coordinates of these corners points were detected in these images.

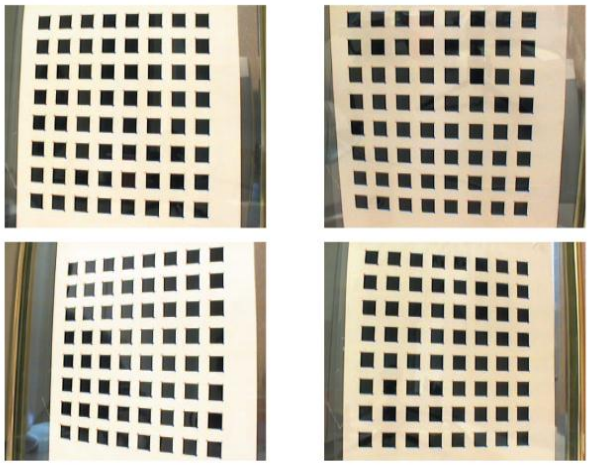

Figure 4.2: Calibrating the target in different orientations (reproduced from (Zhang, 2000)).

The five intrinsic and the six extrinsic parameters were calculated using a closed form solution by initially considering the radial distortion to be zero. The coefficients of radial distortion were determined by least squares methods. The radial distortion parameters were then included in maximum likelihood criteria to refine the estimated camera parameters.

4.4 Camera calibration using a 1D object

Zhang (2004) proposed a technique for camera calibration using a 1D object with three or more collinear points as shown in Figure 4.3. Various experiments were performed to calibrate the camera. It was proved that camera calibration was impossible using a freely moving 1D calibration object consisting of any number of

collinear points. Experiments were also conducted with one end of the calibration object fixed while the other end was allowed to move. Camera calibration was possible using three or more collinear points on the 1D calibration object. Camera calibration accuracy increased with the number of collinear points.

Figure 4.3: 1D calibration object: three beads attached to a stick (reproduced from (Zhang, 2004)).

A closed-form solution was developed by taking six or more observations of the one-dimensional calibration object. A maximum likelihood criterion was used to improve the accuracy of the estimated parameters. This camera calibration technique is suitable for calibrating multiple cameras in scenarios where the calibration object is visible from all the cameras.

5 SELECTING CORRESPONDING LANDMARK POINTS FOR STEREO-PHOTOGRAMMETRY

The accuracy of 3D coordinate determination in stereo-photogrammetry (Chapter 3) depends on selecting identical corresponding landmark points on both the digital images (left and right digital image). Stereo-image matching is of particular importance in stereo-photogrammetry where a small error in selecting identical pixels or landmark points on both the images might result in large errors in the 3D coordinates of the landmark point in the *X*, *Y*, and *Z* directions. It is easy to manually select corresponding landmark points on two digital images if there are unique features available on the ROI. However, it is very difficult to select landmark points on human subjects where landmarks are often poorly defined. Different techniques for corresponding point/patch/feature matching should therefore be used to increase the ease and accuracy of selecting corresponding landmarks on different images.

The techniques are outlined in the flowchart in Figure 5.1. Initially, different matching techniques that are available are described and the matching techniques, which suit the project requirements, are identified and discussed in more detail. The constraints, which enable the search for best matches, are also discussed.

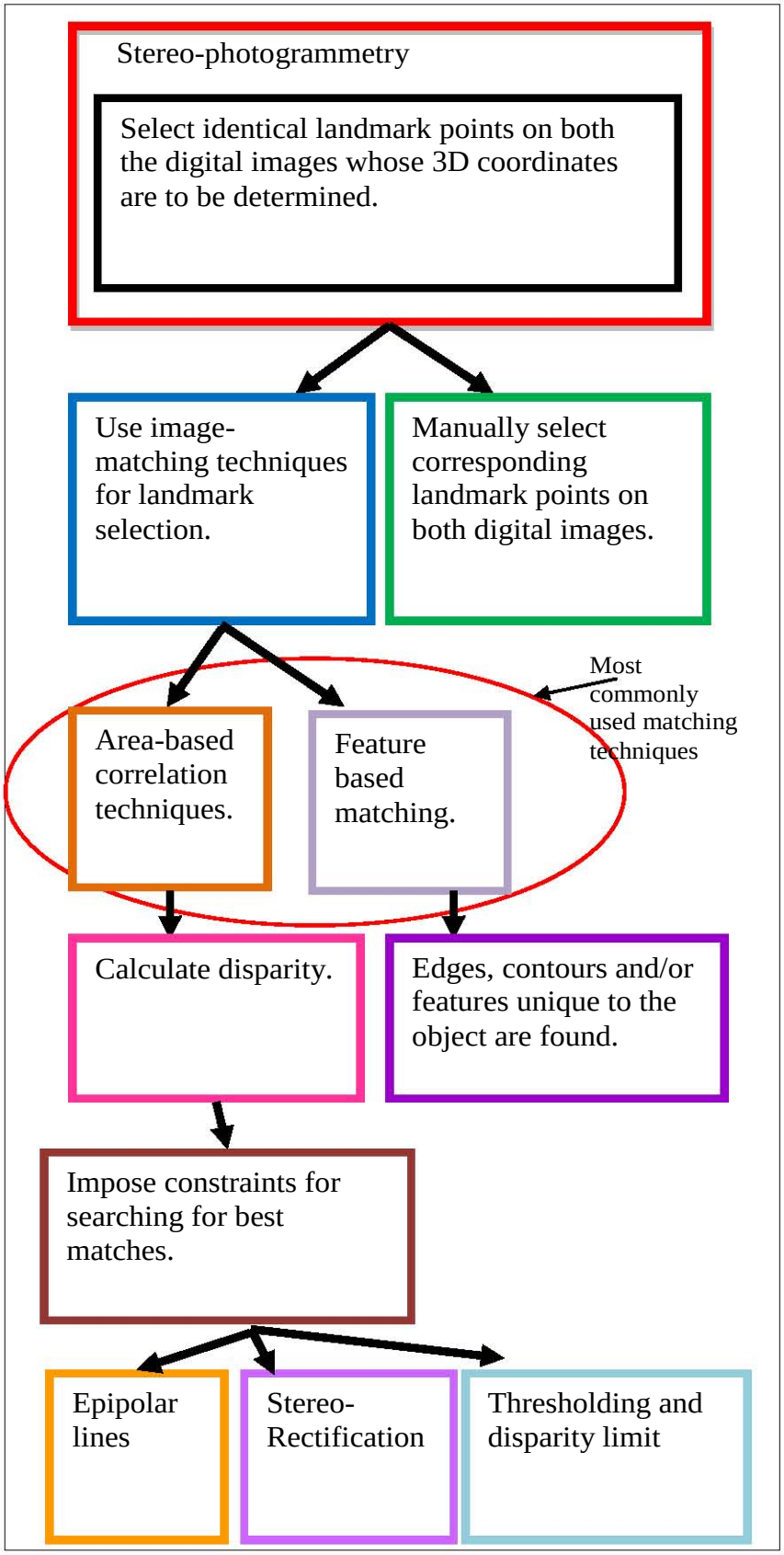

Figure 5.1: Outline of the techniques discussed in this chapter.

5.1 Image matching techniques

Image matching techniques can be area based, transform based, feature based, phase based, hybrid, a combination of different matching techniques, or dynamic (Bennamoun and Mamic, 2002). The discussion that follows assumes that a landmark has been identified in one image, and that a corresponding landmark needs to be searched for in the second image.

Area based image matching

Area based algorithms find matching blocks of image patches in digital images. Either similarity measures are maximised or errors are measured and different techniques are used to minimize the errors. Area based matching may be based on intensities; available techniques include the sum of absolute differences (SAD), the sum of squared distances (SSD) and normalized cross correlation (NCC) (Azad *et al*., 2008). Area based methods are based on in a patch either of the whole image (dense point matching (Fua, 1991)) or of a particular region of the image (sparse point matching (Hannah, 1988)). Other applications of matching techniques include video coding, multi view image generation, camera calibration, 3D reconstruction from stereo-image pairs, object recognition, visual control applications, motion estimation, pose control and many more (Roma *et al* ., 2002).

Transform based image matching (Rank and Census transform)

Marr and Poggio (1979) developed a five-stepped algorithm for stereoscopic matching problem which are as outlined below.

In the first step, the images are filtered at different orientations with bar masks of four sizes. The second step involves localisation of zero crossing in the filtered images, which corresponds to edges, and finding positions of the ends of lines and edges. In the third step, matching is done between pairs of zero crossing or termination of the same sign in the two images. The fourth step involves control of vergence (angle between two cameras) movements by wide masks resulting in the achievement of correspondence of small masks. The fifth step deals with the storage of the correspondences thus achieved in a dynamic buffer.

Zabih and Woodfill (1994) proposed the use of non-parametric local transforms for correlation, based on the relative ordering of local intensity values instead of the intensity values themselves. They found that correlation using this transform can work for outliers and have improved performance near object boundaries when compared with conventional methods such as normalized correlation. It was also observed by them that the non-parametric transforms depend upon the comparative intensities of a pixel versus the pixels in the neighbourhood. They also proposed two non-parametric transform methods. The first method known as rank transform measures local intensity. The second method known as census transform summarizes local image structure.

Feature based image matching

In feature based image matching, unique features such as edges (Hsieh *et al.*, 1997), image patches (Abbasi-Dezfouli and Freeman, 1994), and/or graph structures are matched (Shapiro and Haralick, 1985) in both images. The feature based algorithms

for image patch matching are based on searching predefined features such as edges, contours, unique features invariant to scaling, rotation, translation, illumination changes (Bennamoun and Mamic, 2002; Roma *et al*., 2002).

Hybrid matching techniques

Hybrid matching techniques use a combination matching techniques e.g. areas combined with edge information (Cochran and Medioni, 1992).

Phase based image matching

Phase based image matching techniques use the phase component of the 2D Discrete Fourier Transform (DFT) of images for finding the best match (Miyazawa *et al*., 2005).

Dynamic programming for image matching

Dynamic programming uses dynamic information from images for image matching. Son and Mita (2004) proposed to use the Vitebri algorithm in a stereo correspondence problem. The matching process consisted of matching the right scene to the left scene and again matching the left scene to the right scene. The result of the matching algorithm was selected by comparison between the results of the two matching process. Ohta and Kanade (1983) proposed a stereo-matching technique that uses dynamic programming for obtaining the correspondence between the right and left images. In the method proposed, the intra-scan line and interscan searches proceeded simultaneously. When the digital images are rectified, the epipolar lines become horizontal and the correspondence is searched for along this horizontal scanline (intra-scan line search). Sometimes, there is a mutual dependency between

the horizontal and vertical scan line. The interscan line search tries to find the consistency between the scanlines for finding the best corresponding match. Intille and Bobick (1994) developed a dynamic programming model that finds the correspondence matches and occlusion simultaneously. A data structure called "the disparity space image" (DSI) was defined which modelled the effects of occlusion regions on stereo matching. The corresponding epipolar line from the left and right digital image was used to develop DSI. The right epipolar line was slid on the left image epipolar line pixel by pixel and the difference in the pixel intensities at each pixel was stored in the DSI. When all the pixels were overlapped from the right epipolar line on to the right epipolar line, the DSI was cropped to note the disparity shift.

Different image matching techniques are selected depending on task requirements, trade-off for matching cost, and/or discontinuities in the available data. In addition, the selection of matching technique is affected by illumination variation in the images, partial visibility of the object of interest in one or both images, and occlusion and/or presence of additional objects in one or both images (Roma *et al*., 2002).

Area based and feature based matching techniques are the most commonly used matching techniques and are easy to use and implement. These matching techniques are available in most image processing libraries. The feature based and area based techniques are therefore discussed in detail below.

5.2 Area based correlation techniques

In area based correlation techniques, either a similarity measure is calculated and maximized or an error measure is calculated which is minimized. The correlation technique could be normalized or non-normalized. The normalized correlation functions are independent of constant additive or multiplicative brightness differences between the images. Non-normalized correlation functions are affected by changes in illumination of the scene (Azad *et al*., 2008).

Figure 5.2 maps the area-based correlation techniques described in the remainder of this section.

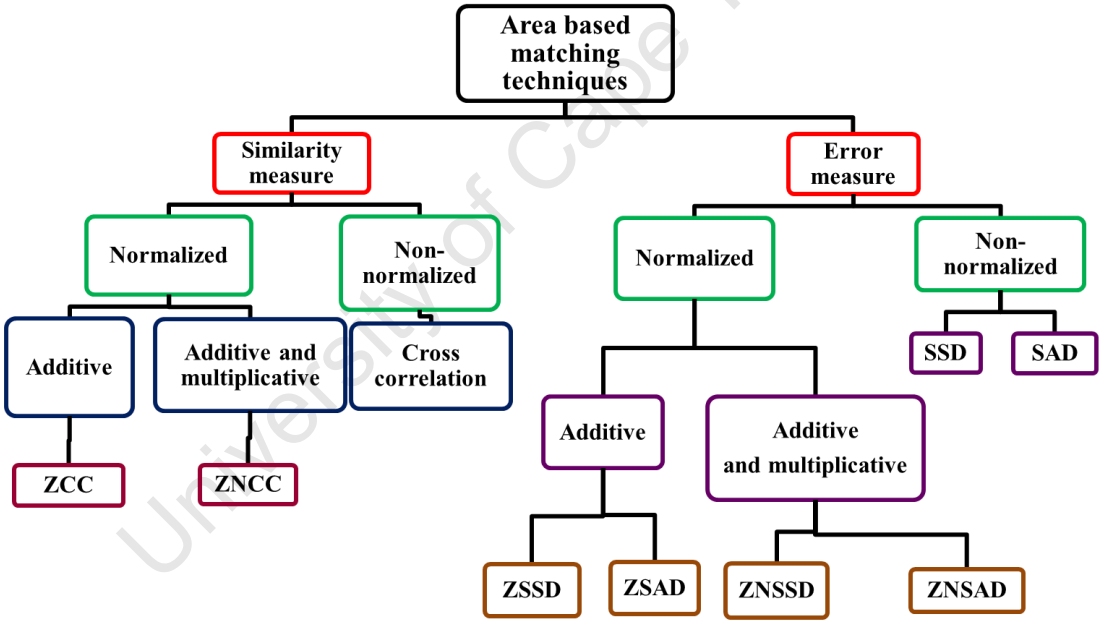

Figure 5.2: Area-based matching techniques (Azad *et al*., 2008).

5.2.1 Non-normalized matching functions

The different non-normalised matching techniques are given below:

- **Cross correlation** defines a similarity measure. The larger the value, the more similar are the images (Azad *et al*., 2008). The function is computationally inexpensive.
- **The sum of squared distances (SSD)** calculates an error measure. If the images for correspondence are identical, the output of this function is zero. The function is reliable if there are similar brightness conditions in both the images; however, the function is sensitive to outliners due to squaring (Azad *et al*., 2008; Banks *et al*., 1997). The function is easy to implement and requires little computational power (Sun and Peleg, 2004). It is sensitive to intensity distortion where the pixel values in one image differ from those in another by a constant offset and/or gain factor (Banks *et al*., 1997).
- **The sum of absolute differences (SAD)** calculates an error measure and is less sensitive to outliners than SSD (Azad *et al*., 2008).

5.2.2 Zero mean matching functions invariant to additive brightness differences

Zero mean matching functions are invariant to constant brightness differences. For example, if one digital image is taken in a bright light and another image is taken in a dim light, the normalized matching functions will give results that are unaffected by the constant brightness differences in the two digital images. The different zero mean "additive" matching techniques are given below:

- **Zero mean cross correlation (ZCC)**: The cross correlation function (Section 5.2.1) is normalized additively (Azad *et al*., 2008). The computational cost increases due to normalization.
- **Zero mean sum of squared differences (ZSSD)**: The SSD function (Section 5.2.1) is normalised additively (Azad *et al*., 2008).
- **Zero mean sum of absolute differences (ZSAD)**: The SAD function (Section 5.2.1) is normalised additively (Azad *et al*., 2008).

5.2.3 Zero mean matching functions invariant to multiplicative brightness differences

Sometimes, pixels of the images differ by a constant multiplicative factor. Zero mean normalized functions are normalized with respect to additive as well as multiplicative brightness differences (Azad *et al*., 2008).

- **Zero mean normalized cross correlation (ZNCC)**: The function cross correlation (Section 5.2.1) is normalized with respect to additive and multiplicative brightness differences (Azad *et al*., 2008).
- **Zero mean normalized sum of squared differences (ZNSSD)** and **Zero mean normalized sum of absolute differences (ZNSAD)**: The functions SSD and SAD (Section 5.2.1) when normalized with respect to additive and multiplicative brightness differences are known as ZNSSD and ZNSAD respectively. These methods are computationally very expensive and require longer time to execute (Azad *et al*., 2008).

5.2.4 Constraints for optimally searching for landmark points

The constraints such as the epipolar line, the disparity limit, and stereo-rectification assist in the search for a landmark point on one digital image that matches a point in the other image (Cyganek and Siebert, 2009). These methods are described below.

5.2.4.1 Disparity and disparity range

Disparity is the measure of the difference between the image coordinates of a landmark point (on two different digital images of the same landmark) when viewed from two different positions by a stereo camera system (Azad *et al*., 2008; Bradski and Kaehler, 2008).

A stereo camera system is illustrated in Figure 5.3. It is assumed that the image planes of the two cameras are coplanar, the images are corrected for distortion, and the principal rays or optical axes are parallel to each other as shown in the Figure 5.3. *'O_l'* and 'O_r' are the centre of projection for the left and the right camera system. '*b*' is the distance between the projection centres for both the camera systems (i.e. between 'O_l' and 'O_r'). The focal length is assumed to be the same for both the cameras and is represented by '*f*'. 'C_{lx}' and 'C_{rx}' are the principal points for the left and right hand cameras and it is assumed that the cameras have been calibrated such that the principal points ('C_{lx}' and 'C_{rx}') have the same pixel coordinates in the left and right hand images (Bradski and Kaehler, 2008).

Point '*P*' is the landmark point in the world coordinate system. 'x^l' and 'x^r' are the horizontal coordinates of the landmark points in left and right digital image respectively. '*r*' is the depth or the normal distance of the landmark point '*P*' from the image plane. The depth '*r*' can be estimated by using the similar triangle theorem (Bradski and Kaehler, 2008).

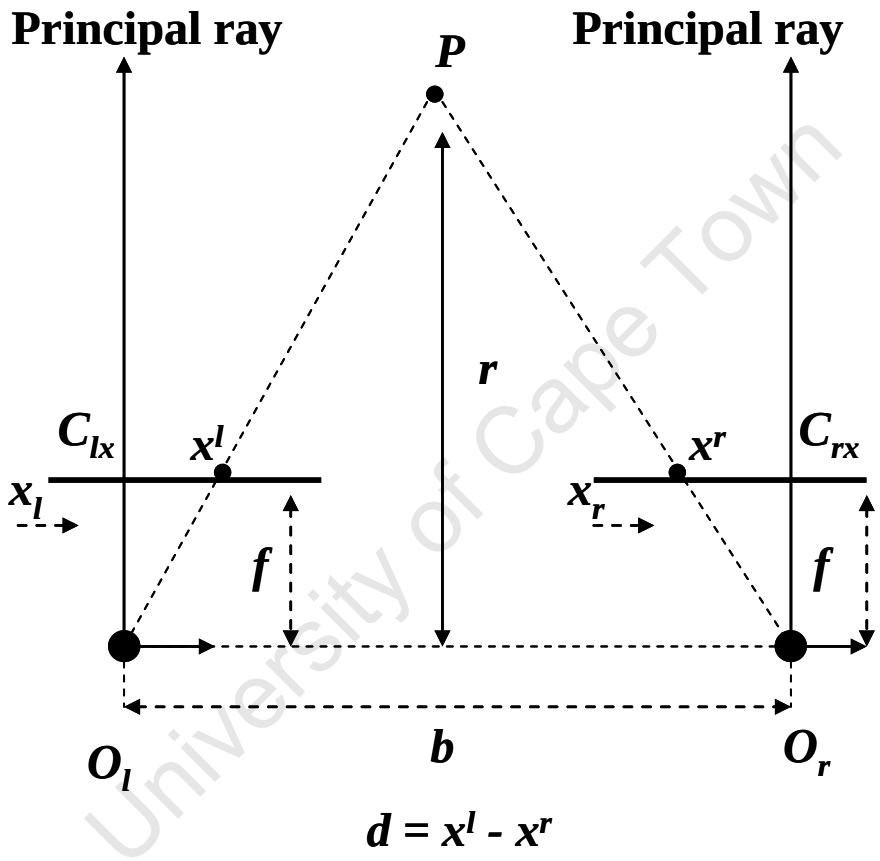

Figure 5.3: Calculation of depth and disparity of the object from the image plane in an undistorted and aligned stereo system. 'O_l' and 'O_r' are the centre of projection for the left and the right camera system. '*b*' is the distance between the projection centres for both the camera systems; '*f*' is the focal length; 'x^l' and 'x^r' are the horizontal coordinates of the landmark points (adapted from (Bradski and Kaehler, 2008)).

The formula for depth is given by (Bradski and Kaehler, 2008):

$$r = bf/d$$ **Equation 5.1**

where

$$d = x^l - x^r$$ **Equation 5.2**

where 'd': disparity, which is the difference between the pixel coordinates of the landmark points when viewed from two different viewpoints.

From Equation 5.1, it is clear that the disparity is inversely proportional to the distance of the object from the image plane as shown in Figure 5.4 (Bradski and Kaehler, 2008). Thus, the farther the object from the cameras, the smaller the disparity range.

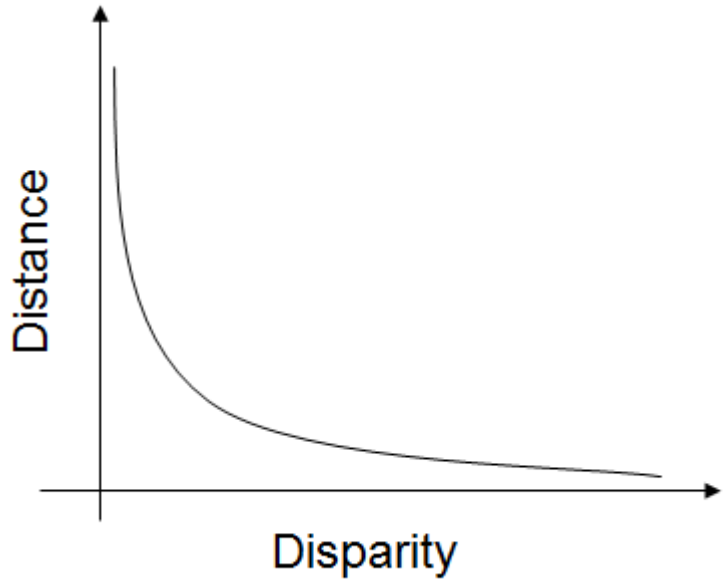

Figure 5.4: Relation between depth and disparity. The distance of the object of interest from the cameras is inversely proportional to the disparity (adapted from (Bradski and Kaehler, 2008)).

For finding correlated points using area based correlation techniques, the relative depth and/or disparity can provide useful information for finding the best match. The search will be limited to the given disparity range. Disparity changes from one landmark to the other and from one image set to the other depending upon the height variance. For example, one landmark point might be farther from the cameras then the other. Therefore, the same disparity value might not be suitable for all the landmark points in a pair of images. Therefore, setting a disparity range covering all the likely disparities between different landmark pairs is useful.

5.2.4.2 Epipolar geometry

Epipolar geometry is used to determine the relationship between the image planes of two or more cameras and 3D points in object space. Epipolar geometry may be used to constrain the search for corresponding landmark points in digital images (Azad *et al.*, 2008).

Consider a 3D point '*P'* in the world coordinate system as shown in Figure 5.5. 'p_l' and 'p_r' are the points on the left and right digital images. The epipolar plane is the plane formed by joining the projection centre 'Z_l', the projection centre 'Z_r', and the 3D point '*P*'. The line Z_l Z_r meets the images at points 'e_l' and 'e_r'. When the epipolar plane intersects with the image plane, epipolar lines are formed. The epipolar line for point '*P*' on the left digital image is 'l_l' and for the same point on the right digital image is 'l_r'. All the points, which correspond to the point 'p_l' in the left digital image, will lie on an epipolar line 'l_r' on the right digital image. The epipolar line looks like a

point when seen from the left camera and it looks like a line from the right camera. Given the camera calibration parameters, 'Z_l', 'Z_r', 'e_l', 'e_r', 'l_l', and 'lr' can be estimated (Azad *et al*., 2008; Bradski and Kaehler, 2008).

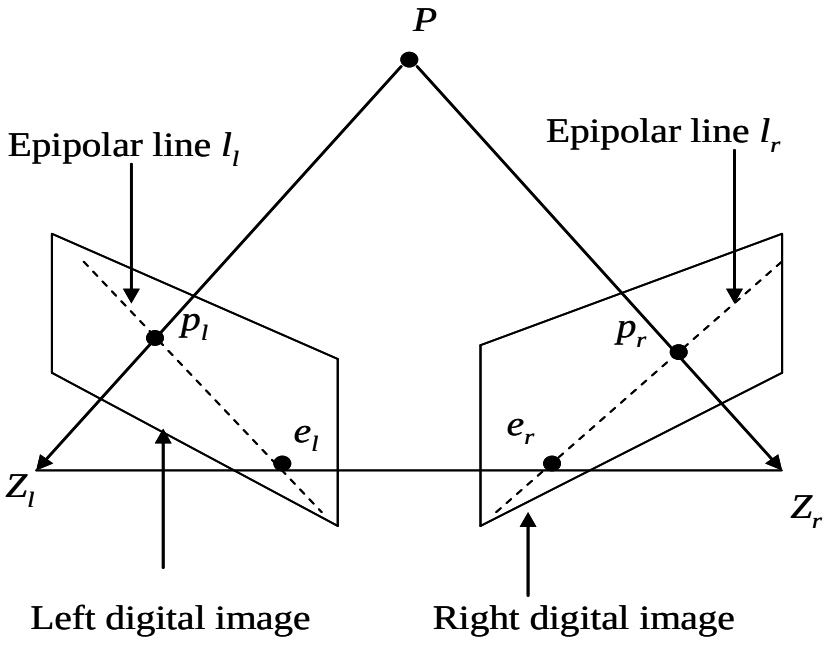

Figure 5.5: Epipolar geometry for the left and right digital image. 'l_l', and 'lr' are the epipolar lines, 'P' is the landmark point in the world coordinate system, 'p_l' and 'p_r' are the landmark points on the left and right digital image (adapted from (Bradski and Kaehler, 2008)).

5.2.4.3 Stereo-rectification

Rectified input digital images can be used to simplify disparity and disparity map estimation (Azad *et al*., 2008). In a real stereo system, the two cameras might not always have exactly coplanar imaging planes i.e. the image planes of the cameras are not parallel to each other due to the relative position of cameras with respect to each other. The goal of the stereo-rectification process is to reproject the image planes of the two cameras so that they lie in the same plane (Bradski and Kaehler, 2008). In

other words, stereo-rectification makes the epipolar lines for both digital images parallel to each other and the corresponding epipolar lines have same *y* coordinate in the image system as shown in Figure 5.6(b) (Bradski and Kaehler, 2008).

Stereo rectification can result in strong distortions in digital images. The distortion depends on the angle between the two cameras (Azad *et al*., 2008).

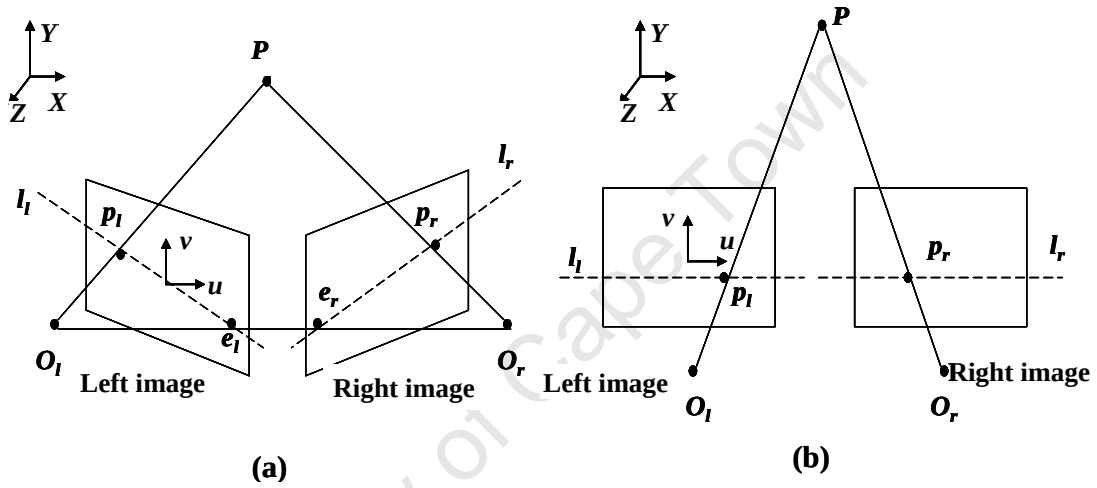

Figure 5.6: Stereo rectification(a) unrectified left and right digital image with non-coplanar image planes; the epipolar lines 'l_l' and 'l_r' are not parallel to each other; (b) rectified left and right digital images with epipolar lines 'l_l' and 'l_r' parallel to each other (adapted from (Belbachir, 2009)).

5.2.5 Implementation of area based correlation

Azad et al (2008) present an area based matching technique as follows. For a clicked point in the left digital image, an image patch of size $(2n + 1)$ x $(2n + 1)$ (where n is greater than or equal to one) is cut out. This image patch could be normalized with respect to additive and/or multiplicative brightness differences depending upon the method selected. Correspondences to this image patch are then searched for in the

right image using any of the area-based correlation techniques. The search space can be reduced by searching for the match along the epipolar line (Section 5.2.4.2). The correspondence is searched for only in a given disparity range. Smaller disparity values represent greater distance from the camera and vice versa (Section 5.2.4.1). The result obtained by this matching technique can be validated by comparing it with a threshold value. If the results are greater than this value, the correspondences are accepted (Azad *et al*., 2008).

5.3 Feature detection (line and circle detection)

Feature based methods are well-established image matching techniques in stereovision. Features such as retro-reflective markers can be used as a unique feature on human subjects in order to make the selection of corresponding landmarks easier.

The Hough transform detects edges as well as specific features (such as lines, ellipses, and /or circles) in digital images (Nixon and Aguado, 2008). The Hough line and circle transforms for detecting lines and circular markers are discussed below.

5.3.1 Polar Hough line transforms

The Hough line transform may be used to detect lines in a chessboard pattern while calibrating cameras (De la Escalera and Armingol, 2010). The Hough transform for a line in Cartesian and polar coordinates is given below (Nixon and Aguado, 2008).

In Cartesian coordinate system, a line formed by collinear points with coordinates (*x*, *y*), slope '*m*', and intercept '*c*' is given by:

$$y = mx + c \qquad \textbf{Equation 5.3}$$

Equation 5.3 can be written in homogenous form as:

$$Ay + Bx + 1 = 0 \qquad \textbf{Equation 5.4}$$

where $A = \frac{-1}{c}$ and $B = \frac{m}{c}$

Thus, a line can be defined by a pair of values (A, B). Consider that the set of collinear points $\{(x_i, y_i)\}$ define the line (A, B), then Equation 5.4 can be rewritten as:

$$Ay_i + Bx_i + 1 = 0 \qquad \textbf{Equation 5.5}$$

Equation 5.3 can be written in Cartesian coordinates as:

$$c = -x_i m + y_i \qquad \textbf{Equation 5.6}$$

The relationship between a point (x_i, y_i) and the line given by Equation 5.6 is shown in Figure 5.7. The points (x_i, y_i) and (x_j, y_j) in Figure 5.7(a) define the lines L_i and L_j in dual or accumulator space in Figure 5.7(b).

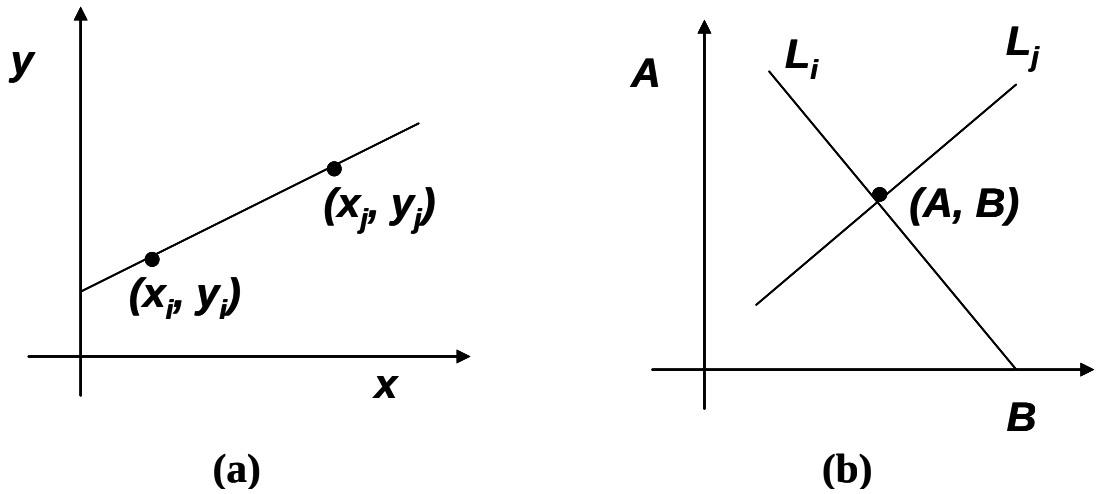

Figure 5.7: A line in Cartesian and dual space: (a) image with a line formed by two points (x_i, y_i) and (x_j, y_j) in the Cartesian coordinate system, (b) two lines 'L_i' and 'L_j' formed by the two points (x_i, y_i) and (x_j, y_j) in dual space. The point (A, B) is the point where the two lines meet. The parameters 'A' and 'B' can be used to define the equation of the line from Equation 5.5 (adapted from (Nixon and Aguado, 2008)).

Accumulator or dual space is an array having elements equal to that of all the collinear points defining the line. All the collinear points defining a line will have the same concurrent point (A, B) as shown in Figure 5.7(b). The axis in the dual space represents parameters of the line $((A, B) or (m, c))$. The Hough transform uses an accumulator array that stores the votes, which is incremented by one for each point lying on the line as shown in Figure 5.8 (Yu, 2010). In other words, if four points lie on a line, then the accumulator array indicating that line will have a value 4 (the value is in the accumulator array is incremented by one for each point lying on the line). All the points lying on the dual lines are traced and the accumulator array is incremented for each point on the dual line. The accumulator array is then searched for the maximum count to find the parameters of the line. As shown in Figure 5.7(b), the two lines meet at (A, B) which can be used to determine the equation of the line

formed by the collinear points (Equation 5.3 or Equation 5.4). Thus, the line can be detected in the digital image using the Hough transform.

In Cartesian parameterization, the slope m varies from zero to infinity from horizontal to vertical lines. The accumulator array might produce bias errors as m takes infinite values. The problem of Cartesian parameterization in the Hough transform is solved by using the parameterized Hough transform or foot-of normal parameterization given below (Nixon and Aguado, 2008). In the parameterized Hough transform, polar coordinates are used instead of Cartesian coordinates for indicating the parameters of a line.

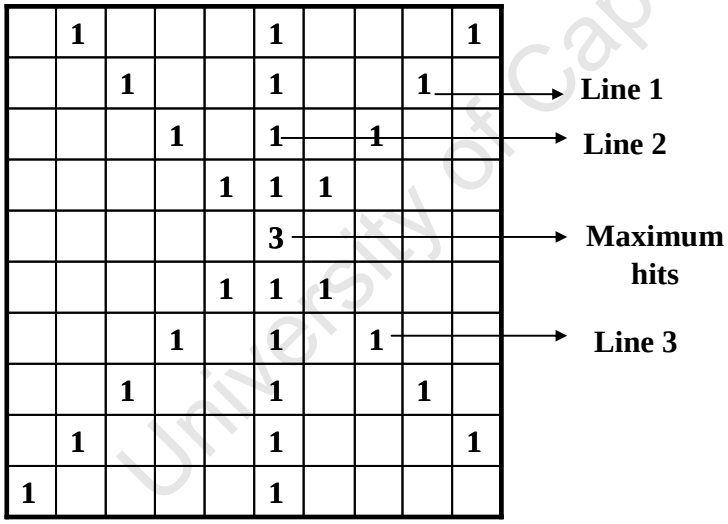

Figure 5.8: Accumulator array. The cell having maximum count gives the parameters of the line (*m*, *c*) or (*A*, *B*) (adapted from (Yu, 2010)).

In the parameterized Hough transform, a perpendicular line is drawn from a point (x, y) to the origin as shown in Figure 5.9. The point where this line intersects this perpendicular line in the image is given by:

$$\rho = x\cos(\theta) + y\sin(\theta) \qquad \textbf{Equation 5.7}$$

Where θ is the angle of the perpendicular line from origin to the line in an image and ρ is the length between the origin and the point of intersection of the perpendicular line from origin to the line in an image.

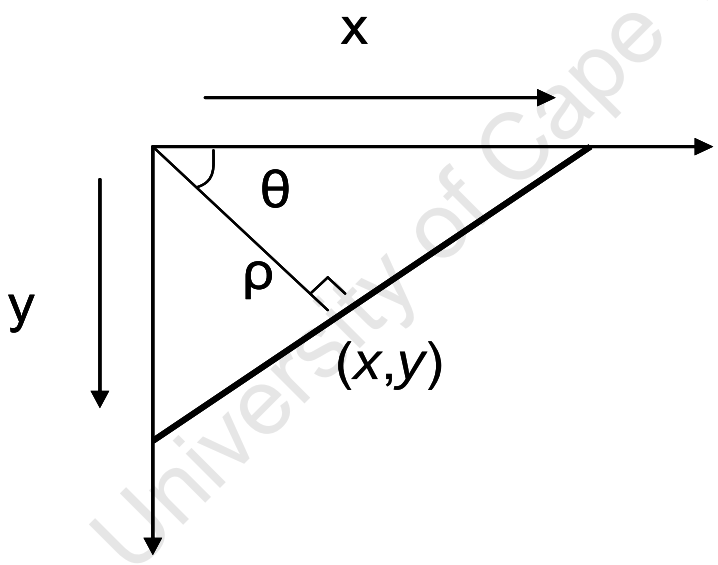

Figure 5.9: Representation of a line in polar coordinates. θ is the angle of the perpendicular line from the origin to the line and ρ is the perpendicular distance of the intersection of both the lines from origin (adapted from (Nixon and Aguado, 2008)).

In the polar Hough transform, a point maps to a curve in accumulator space as shown in Figure 5.10(a). The accumulator spaces for one, two, and three points are shown in

Figure 5.10 (a), (b), and (c). The point where the three curves meet in the accumulator space gives parameters that can be used to derive the equation of a line.

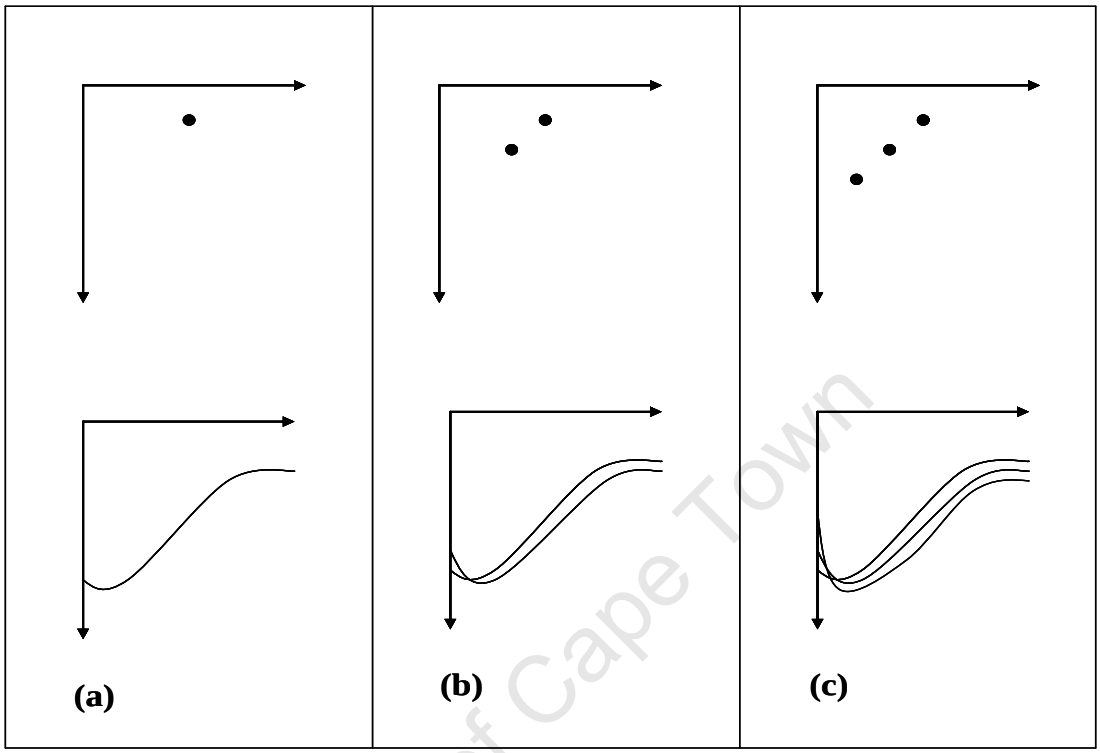

Figure 5.10: Images and accumulator space of the polar Hough transform. The top row indicates a point in polar coordinates and the bottom row indicates the point as a curve in the accumulator space for (a) one landmark point, (b) two landmark points, (c) three landmark points. The coordinates of the point of intersection of the curves in the accumulator space gives the parameters of the line (adapted from (Nixon and Aguado, 2008)).

5.3.2 Hough circle transform

Circular markers might be used as a unique feature in digital images to facilitate the selection of corresponding landmarks. The Hough line transform (Section 5.3.1) can be extended to find circles or circular markers in digital images (Nixon and Aguado, 2008). The method, taken from (Nixon and Aguado, 2008), is given below.

The graph of a circle with origin (a,b), locus of points (x_i, y_i), and radius R in explicit form is shown in Figure 5.11 and the equation of a circle is given by:

$$(x_i - a)^2 + (y_i - b)^2 = R^2 \qquad \textbf{Equation 5.8}$$

Or in polar coordinates, the equation of the circle is given by:

$$x = a + R\cos(\theta) \qquad \textbf{Equation 5.9}$$

$$y = b + R\sin(\theta) \qquad \textbf{Equation 5.10}$$

where θ varies from 0 to 360 degrees and the points (x, y) trace the perimeter of the circle. If the image has many points lying on the perimeter of the circle, then to describe a circle, (a,b,R) need to be searched for in accumulator space. Therefore, the parameter space is 3D for the Hough circle transform instead of 2D as for the line transform. If the radius of the circle is known, then the 3D accumulator space can be reduced to 2D accumulator space (Yu, 2010).

Each edge point on the circumference of the circle represents a circle with edge point as the centre and R as radius in accumulator space. Figure 5.12 (a) shows three edge points on the circle and Figure 5.12(b) shows three circles formed by three edge points in accumulator space. The accumulator array (similar to Figure 5.8) is incremented for each edge point and the array element having the highest value gives the parameters of the circle.

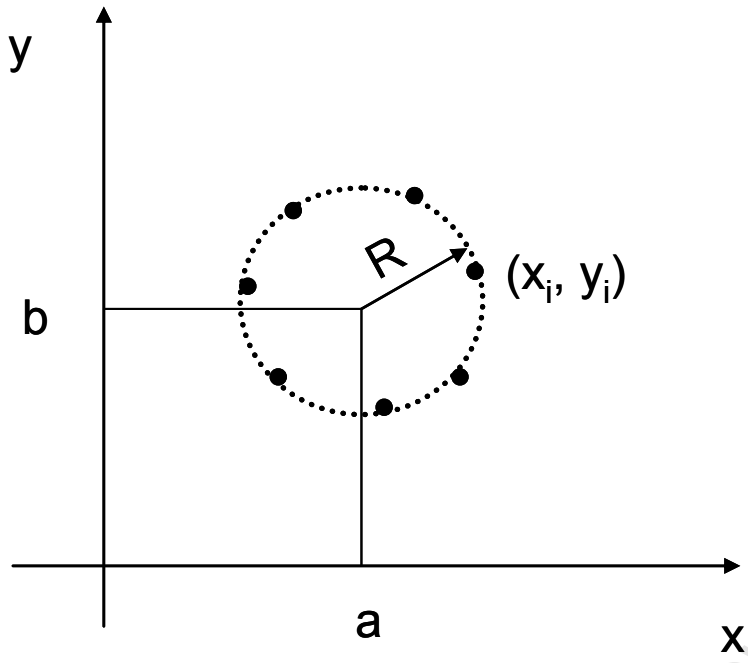

Figure 5.11: Representation of circle in explicit form. (*a*, *b*) represent the centre and '*R*' is the radius of the circle (adapted from (Nixon and Aguado, 2008)).

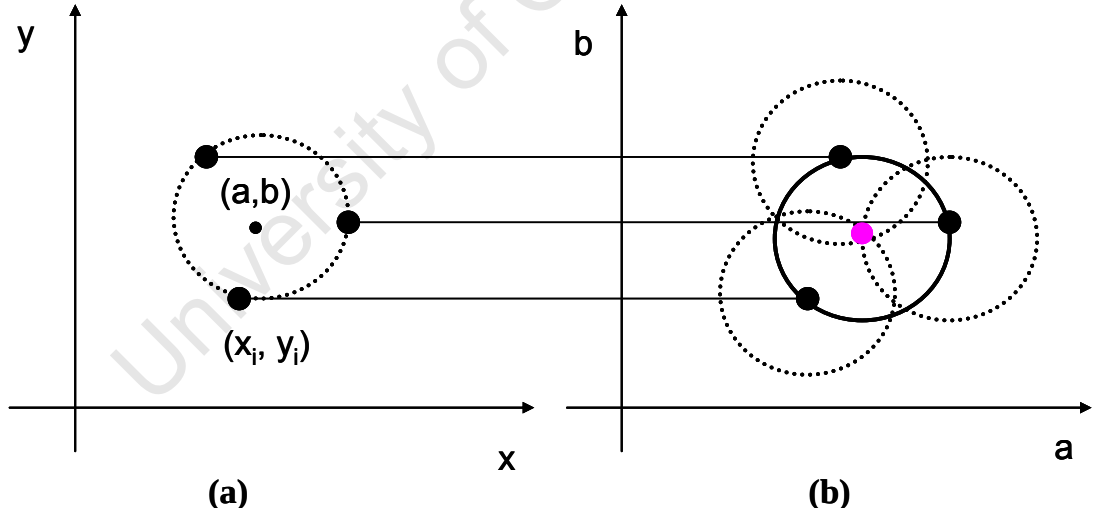

Figure 5.12: Hough circle transform: (a) image containing three edge points on the perimeter of the circle; (b) each edge point represents a circle in accumulator space and the intersection of all the circles (pink dot) gives the parameters for the equation of a circle (adapted from (Nixon and Aguado, 2008)).

Detecting the edges in images is a general pre-processing step in searching for circular markers. Once the edges are determined, the Hough circle transform may be used to detect circular markers (Azad *et al*., 2008).

A commonly used edge detector is the Canny edge detector. It detects the edges in the digital images with high accuracy, reduces noise by applying Gaussian smooth filter, supresses the non-maximum edge points, and thresholds with hysteresis to connect the edge points. The non-maximum suppression results in thin lines of edge points, by retaining all the points at the top of a ridge of edge data (Nixon and Aguado, 2008). Hysteresis thresholding (Figure 5.13) starts when an edge point obtained from non-maximal suppression exceeds an upper threshold. This edge point is termed as a seed point and is set to white (or is labelled as an edge point). This edge point forms the first point of a line of edge points. The points in the neighbour of this edge point are then checked to determine if they exceed lower threshold. All the neighbouring edge points exceeding the lower threshold are also labelled as edge points or are set to white as shown in Table 5.1 (Nixon and Aguado, 2008). The search terminates at points where no neighbouring edge points are above the lower threshold.

The Canny edge detector decreases the number of edge points to be examined in the search for circular markers on the object.

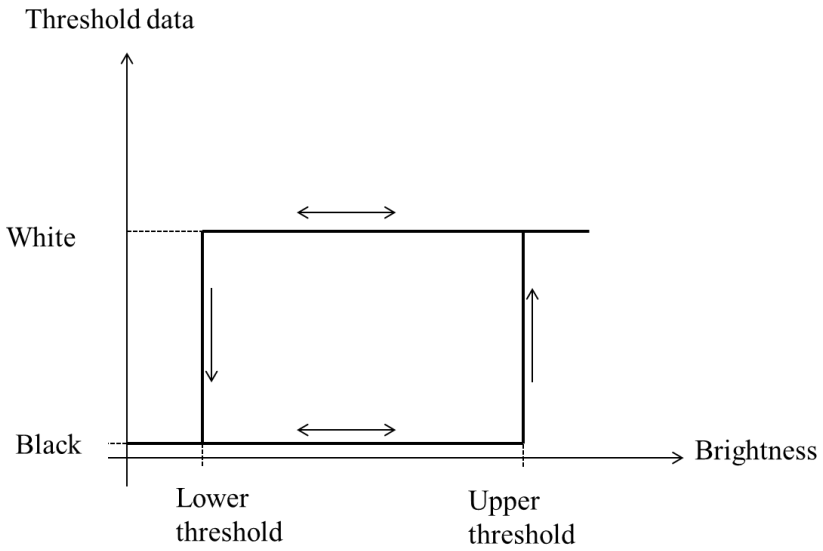

Figure 5.13: Hysteresis thresholding. Points are set to white if the upper threshold is exceeded and set to black if the lower threshold is reached (adapted from (Nixon and Aguado, 2008)).

Table 5.1: All the points in the neighbourhood of a seed point (at the centre) are checked if they exceed a lower threshold. If the neighbouring points exceed the lower threshold, they are labelled as seed points and the search for edge points continues (adapted from (Nixon and Aguado, 2008)).

Neighbouring points ≥ lower threshold	Neighbouring points ≥ lower threshold	Neighbouring points ≥ lower threshold
Neighbouring points ≥ lower threshold	Seed point ≥ Upper threshold	Neighbouring points ≥ lower threshold
Neighbouring points ≥ lower threshold	Neighbouring points ≥ lower threshold	Neighbouring points ≥ lower threshold

6 PRELIMINARY TEST USING AVAILABLE EQUIPMENT

An experiment was conducted to test the accuracy of the cameras and lenses to be used in the project, in order to determine their suitability for the proposed solution. Preliminary testing using stereo-photogrammetry was done with the help of available equipment (a calibration frame) and software (Australis) used for research by Mutsvangwa et al (2009) and Mutsvangwa (2009).

6.1 Cameras and Lenses

The specifications of the lenses and cameras used in the project are given in Table 6.1 and Table 6.2.

Table 6.1: TAMARON CCTV (CS-mount) Lens specifications (http://www.tamron.co.jp/en/data/cctv/13vm2812as2.html)

Model		13VM2812AS II
Imager Size		1/3
Focal Length		2.8-12 mm
Aperture range		1.4-Close
Mount Type		CS
Angle of View (Horizontal X Vertical)	Wide	97.4° X 72.5°
	Tele	24.1° X 18.1°
Operation	Focus	Manual w/Lock
	Zoom	Manual w/Lock
	Iris	Manual w/Lock

Table 6.2: FIREFLY MV CS-mount camera specifications
(http://www.ptgrey.com/products/fireflymv/fireflymv.pdf)

Specification	FFMV-03MTM/C (BW or Color)
Image Sensor Type	1/3" Progressive scan CMOS
Maximum Resolution	752 (H) X 480 (V)
Pixel Size	6.0 µm X 6.0 µm
Video Data Output	8 and 16-bit digital data
Image Data Formats	Y8, Y16(monochrome), 8-bit, and 16-bit raw Bayer data (color models)
Digital Interface	6-pin IEEE 1394a for camera control, video data, power

6.2 Calibration frame

Mutsvangwa (2009) used the 3D calibration frame shown in Figure 6.1 to determine 3D landmark coordinates. The calibration structure consists of a backbone on which the control frame is clamped. The calibration frame has 91 retro-reflective markers and the diameter of each of the markers is 5mm.

6.3 Software

The Australis software Version 6.06 (Photometrix Pty. Ltd., Victoria, Australia) is software designed to perform highly automated image-based 3D coordinate measurement from digital images. It was used to calibrate the cameras and to calculate the three-dimensional coordinates of the control markers on the calibration frame.

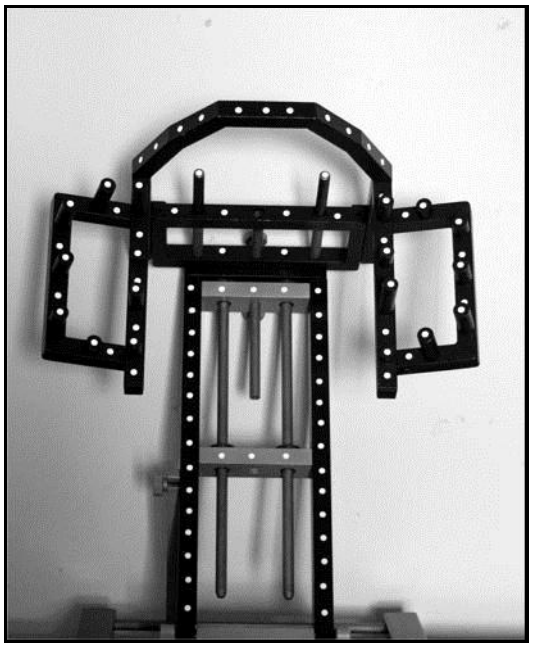

Figure 6.1: 3D calibration frame with 91 retro-reflective circular markers (Mutsvangwa, 2009).

6.4 Camera calibration

Eight images of the calibration frame taken from different orientations were used to calibrate the camera (Figure 6.2). All the reference markers (91) on the calibration frame shown in Figure 6.1 were used to calibrate the camera. The 3D coordinates of the centres of the circular grid markers were known (accuracy 0.1 mm). The calibration frame retro-reflective marker centres were selected by mouse click in each image. The Australis software computed the 3D coordinates of the centre of the circular markers selected on the screen. Bundle adjustment (Section 3.6) was used to determine their 3D coordinates.

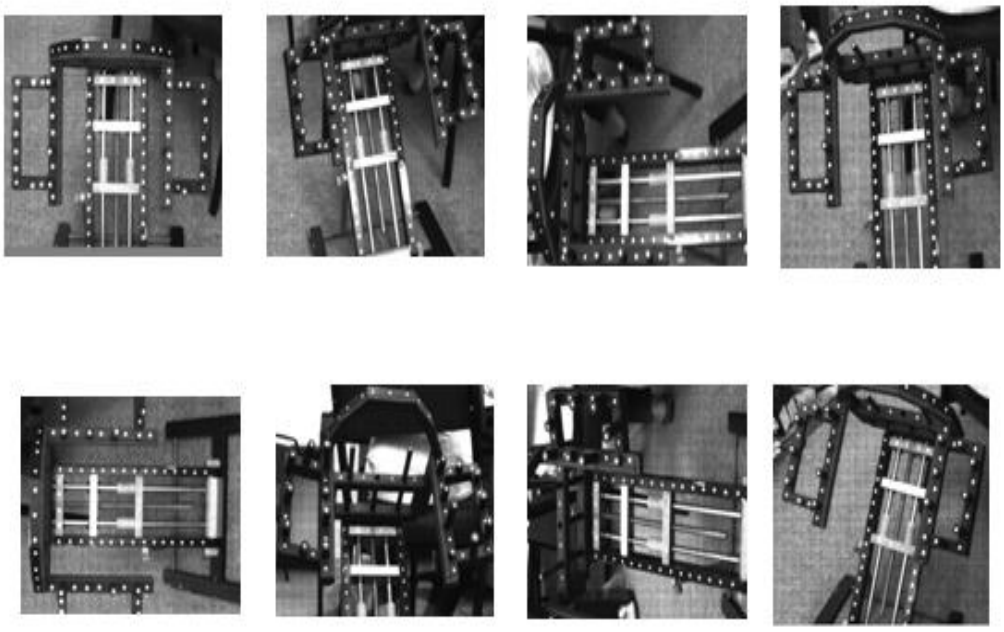

Figure 6.2: Camera calibration using Australis. The calibration frame was imaged in different orientations. Eight images of the calibration frame were considered for calibration. All the reference markers (white dots) were used to calibrate cameras.

6.5 Finding 3D coordinates

The 3D coordinates of all 91 retro-reflective markers were already known. The accuracy of the application was tested by finding the 3D coordinates of 23 check markers shown in Figure 6.4 using 47 control markers (Figure 6.3) on three test images of the calibration frame. The Australis software using bundle adjustment was used to find 3D coordinates of the 23 check markers. The 3D coordinates thus obtained were compared with the actual 3D coordinates. Mean error and standard deviation for the actual 3D coordinates of the markers and the 3D coordinates obtained using the Firefly CS mount camera in *X*, *Y* and, *Z* directions are shown in Table 6.3.

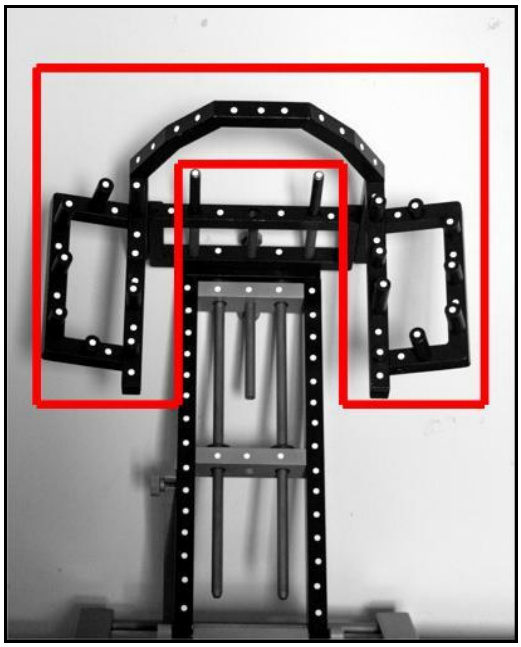

Figure 6.3: Control markers on the calibration frame.

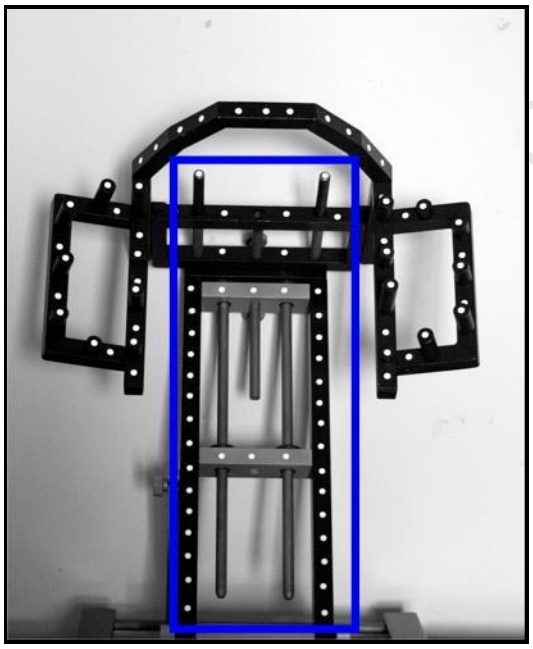

Figure 6.4: Check markers on the calibration frame.

Table 6.3: Results for 3D measurement by stereo-photogrammetry using digital images taken using CS-mount cameras.

Plane	Mean error(in mm)	Standard deviation (in mm)
X-plane	0.1	0.5
Y-plane	0.7	0.5
Z-plane	0.3	0.2

6.6 Conclusion

The errors in the X, Y and Z directions were between 0.1 to 0.7 mm. The results were comparable to those obtained by Mutsvangwa (2009) using high-resolution cameras (Canon 350D Digital SLR with 18-55 mm variable-lens). Mutsvangwa (2009) obtained 0.3 mm errors in the X and Y directions and 0.1 to approximately 2 mm in the Z direction. As discussed in Section 1.2, errors up to 13 mm are permissible in locating the ROI on patient's body. From the results obtained, the camera and lens are suitable for the proposed solution.

For the current project, a system is required which should be able to capture digital images of the ROI and use these to determine the 3D coordinates of and distances between landmark points to define the ROI prior to scanning. Processing time should be short as the patient might change position before scanning. The Australis software internally uses bundle adjustment for finding 3D coordinates, which is computationally expensive and requires initial approximation (Section 3.6). The calibration frame is cumbersome and unsuitable for use with the X-ray imaging system. A new stereo-photogrammetric system was designed that is better suited to

the imaging system, but which uses the cameras tested in this one; it is described in the next chapter.

7 DESIGN OF THE STEREO-PHOTOGRAMETRIC SYSTEM

7.1 Scanning an ROI

The aim of the project is to locate an ROI on a patient's body prior to X-ray imaging. Stereo-photogrammetry (Chapter 3) is used to determine 3D coordinates of the landmark points using two digital images of the object of interest taken by two different cameras. The application is developed and tested on the Lodox Statscan machine (Figure 1.1). The steps involved in scanning a ROI on the Statscan machine prior to scanning using stereo-photogrammetry are given below:

1) Two cameras, mounted on the C-arm, are calibrated using a chessboard pattern.

2) Digital images of the patient are taken by the pair of cameras on either side of the C-arm.

3) A graphical user interface (GUI) shows the digital images of the ROI taken by the left and right cameras to the operator (the details of the GUI are given in Chapter 8). The operator selects corresponding landmark points defining the ROI on both the digital images (either manually or by using matching techniques). The camera calibration parameters along with the 2D coordinates of the landmark points are triangulated. The output is the 3D coordinates of the landmark points selected.

4) The length, breadth, and the thickness of the ROI are calculated using the 3D coordinates obtained in step 3. In addition, the horizontal distance of the cameras from the ROI is determined based on the camera calibration parameters and the

marked landmark points. This information is used to find the position of the C-arm relative to the ROI.

5) The length and the breadth define the ROI to be scanned and the thickness may be used to select the kilovolt (kV) and milliampere (mA) settings, which determine image quality and radiation exposure to the patient.

6) The C-arm is moved to the start of the ROI prior to scanning.

7) The C-arm travels the length of the ROI while the X-ray image is obtained. The C-arm is stopped once the length of the ROI is covered.

8) The result is an X-ray image of the ROI.

7.2 Camera mounting unit on the Lodox Statscan machine

The cameras were mounted on the C-arm of the Lodox Statscan machine. Two pairs of cameras are used on either side of the C-arm. Depending on the position of the ROI, a particular pair of cameras will take digital images of the ROI. The camera-mounting unit (Figure 7.1) was designed in such a manner that it does not affect normal operation of the Lodox machine. The distance between the two cameras can be adjusted. In addition, the distance between the C-arm and the cameras is also adjustable. The cameras are able to move in different directions and angles. These adjustments may be made in order to optimally cover the ROI with both cameras.

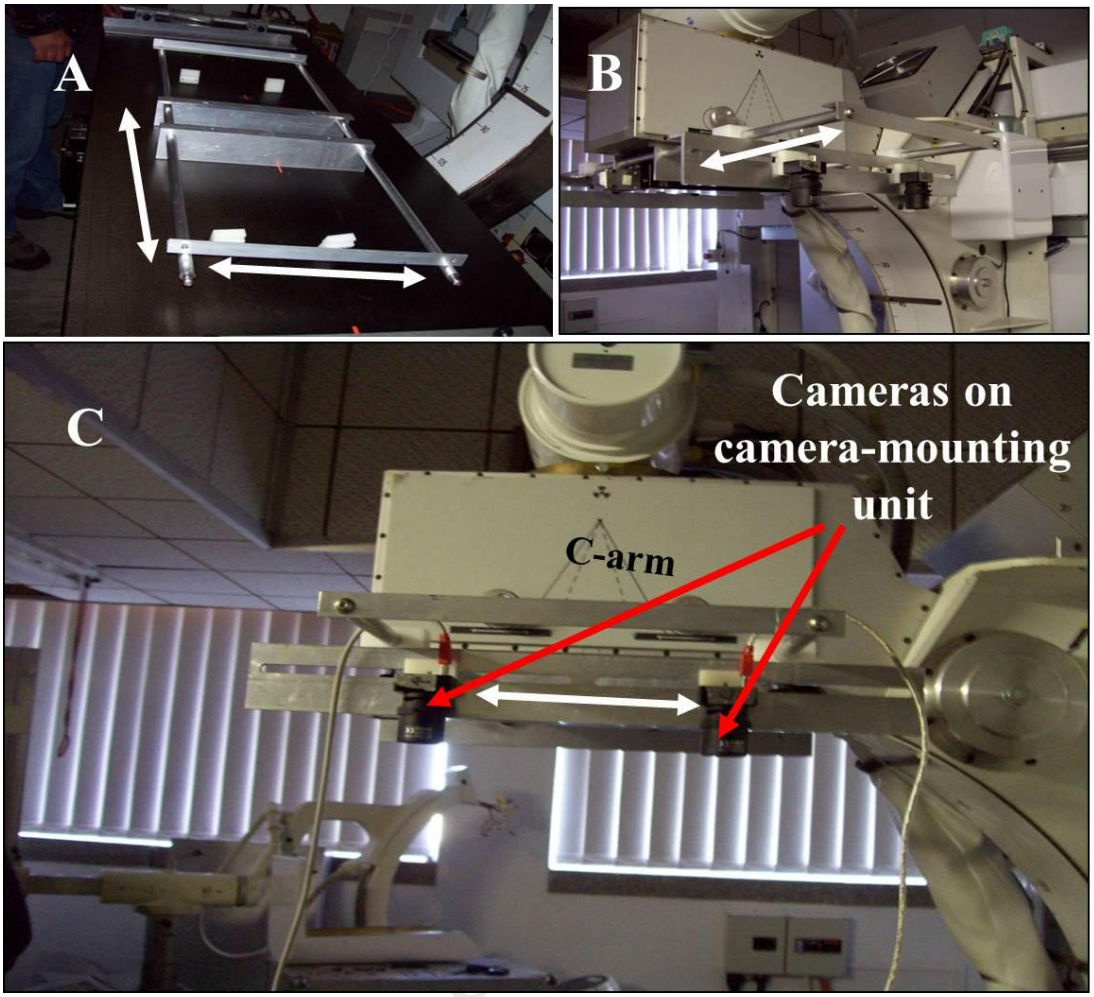

Figure 7.1: Camera mounting unit on the C-arm of the Lodox Statscan machine. A shows the mounting frame; B and C show the cameras mounted on the C-arm. White arrows in A, B, and C indicate that the distance can be adjusted in the arrow direction. The red arrows in 'C' indicate the two cameras.

7.3 Software

The Integrating Vision Toolkit (IVT; Version 1.3.6) (Azad *et al.*, 2008) library and built in applications were used for calibrating two cameras at a time, for capturing digital images simultaneously, and for 3D coordinate measurement. IVT is an image processing library freely available under the GNU license. It has a camera interface and can support openCV, Qt libraries and drivers for firewire cameras. Microsoft

visual C++ 2008 Express edition was used as an integrated development environment (IDE).

7.4 Steps involved in stereo-photogrammetry

The steps involved in finding the 3D coordinates of the landmark points using stereo-photogrammetry are given in Figure 7.2.

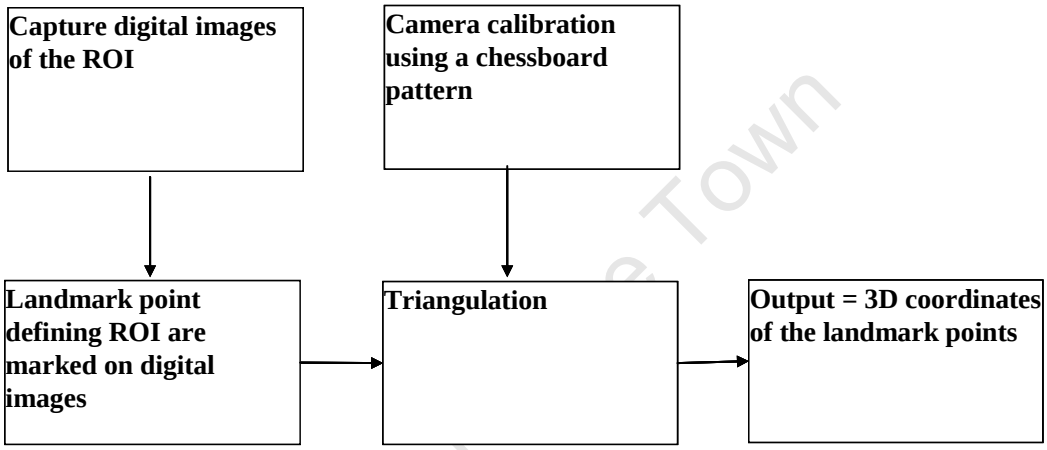

Figure 7.2: Steps for stereo-photogrammetric determination of 3D landmark coordinates.

7.5 Camera calibration and synchronisation

The cameras were calibrated using the chessboard pattern shown in Figure 7.3 as described in Section 4.3. The 3D coordinates of markers on the planar (or 2D) object such as a chessboard need not be known; a 2D planar calibration object is therefore easier to implement than a 3D calibration frame (Section 6.6). Good accuracy can be achieved using this technique and the set-up is very simple (Section 4.3).

The camera calibration application available on IVT (Azad *et al*., 2008) was used for calibrating the cameras. OpenCV libraries were internally used in IVT's camera calibration application for calibrating cameras. The details of the implemented process in IVT for camera calibration may be found out in the books by Azad et al (2008) and Bradski and Kaehler (2008). Four intrinsic parameters (focal length and principal point in x and y directions), five distortion parameters (three radial and two tangential), and six extrinsic parameters (three rotation and three translation) were determined. The intrinsic and extrinsic parameters were determined by Zhang's (2000) method (Section 4.3) and Brown's (1971) method (Section 3.4) was used for distortion correction using the OpenCV libraries (Bradski and Kaehler, 2008).

The theory given in (Bradski and Kaehler, 2008) was used to determine the number of rows and columns for the chessboard frame and the number of digital images of the chessboard pattern to be used. Consider that there are N corners in the chessboard pattern and K images of the calibration frame. Therefore, there are $2NK$ constraints. Considering the distortion to be zero, ten parameters should be estimated (four intrinsic and $6K$ extrinsic parameters for all the orientations). Solving for these parameters require $2NK \geq 6K + 4$ or $(N - 3)\ K \geq 2$. If $N = 5$, ideally K=1 image should solve all the parameters for camera calibration. However, a single image of the chessboard pattern would not be able to solve for all the parameters (there are 10 parameters and from one image, only eight parameters could be obtained (Bradski and Kaehler, 2008)). Therefore, a minimum of two views of a 3x3 chessboard pattern should be used for camera calibration (Bradski and Kaehler, 2008).

However, for high quality results, a nine columns by seven rows chessboard pattern (Bradski and Kaehler, 2008) shown in Figure 7.3 was used for camera calibration. The focus, iris, and field of view were adjusted manually for both the cameras (Table 6.1). The calibration frame was imaged in different orientations. The cameras simultaneously captured digital images of the calibration frame twenty times. In each of the frames captured, the corners of the chessboard pattern (vertical and horizontal lines) were detected by the IVT's camera calibration application and shown on the image. The implicit and explicit parameters of the camera calibration were stored in a text file once the calibration process was complete.

Once fixed, the focus, iris, field of view, and the position of the cameras should not be changed. The cameras should be calibrated again if these parameters or the camera positions changed.

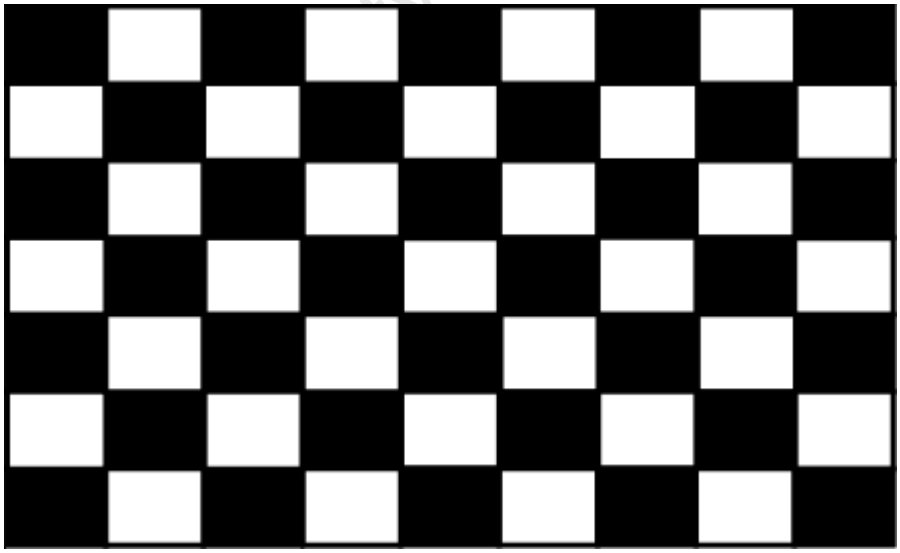

Figure 7.3: Seven rows by nine columns chess board pattern for calibrating two cameras at a time.

7.6 Matching for selecting corresponding landmark points

The accuracy of 3D coordinate measurement depends on selecting accurate corresponding landmark points on both digital images. As an alternative to manual selection in both images, automated matching in the second image after manual selection of landmarks in the first image is used.

The different area based matching techniques are given in Section 5.2. SSD and SAD are computationally simple and work well if the lighting conditions are the same for both the right and the left digital image (Braillon *et al*., 2006). However, the methods fail if there are additive or multiplicative brightness differences between the two images (Azad *et al*., 2008).

For the proposed solution, an additional source of light is not used, so there might be uneven lighting conditions in the room. There might be a slight variation in aperture, iris and/or focal length of the two cameras as these parameters are adjusted manually. These variations may affect matching to find the best match. Therefore, matching techniques independent of brightness differences should be chosen. ZNCC maximizes the similarity measure and is invariant to multiplicative and additive brightness variations in the image. ZSAD minimizes the error measure and is independent of additive brightness variations. The ZSAD function is less sensitive to outliers when compared with other correlation techniques (Azad *et al*., 2008). ZNCC and ZSAD are computationally expensive but in this case, the landmark points selected are very few in number and they are optimally searched for along the epipolar line to reduce the search area and computational time.

ZNCC and ZSAD are therefore selected for the proposed solution. Research has shown that both methods give good matching results (Banks and Corke; 2001; Sun and Peleg, 2004).

In addition to the above mentioned correlation techniques, circular retro-reflective markers (four to six in number) of 32 millimeter diameter are also used as distinct features on the body. The centers of these markers are determined automatically by using the Hough circle transform on both the digital images (Section 5.3.2).

Before applying the area based matching techniques (ZNCC and ZSAD), both digital images are corrected for distortion (to reduce the lens distortion if present) as distortion might affect camera calibration and 3D coordinate measurement of landmark points. The ZNCC and ZSAD matching algorithms in IVT use 640 X 480 resolution, gray scale (8 bit) digital images. Therefore, the 24 bit digital images are converted into 8 bit gray scale images.

The digital images are not rectified, in order to retain the geometrical conditions of the digital images. In the rectification process, the epipolar lines (Figure 5.6) are made parallel to each other (Bradski and Kaehler, 2008) which might result in strong distortions depending on the angle of the cameras to each other (Azad *et al.*, 2008).

7.7 Stereo-triangulation

A 2D calibration frame with a chessboard pattern was used for camera calibration (Section 7.5). The 3D coordinates of landmark points were therefore calculated using the method given in Section 3.5, using the libraries and applications available in IVT. The distance between the two landmark points was calculated to validate the 3D coordinates obtained as the actual 3D coordinates on the objects were not available.

The summary of the implemented process using the IVT image processing library's functions and the sample applications is given in Figure 7.4. A graphical user interface was developed to integrate all the processes into a single product. The details of the product are given in the Chapter 8.

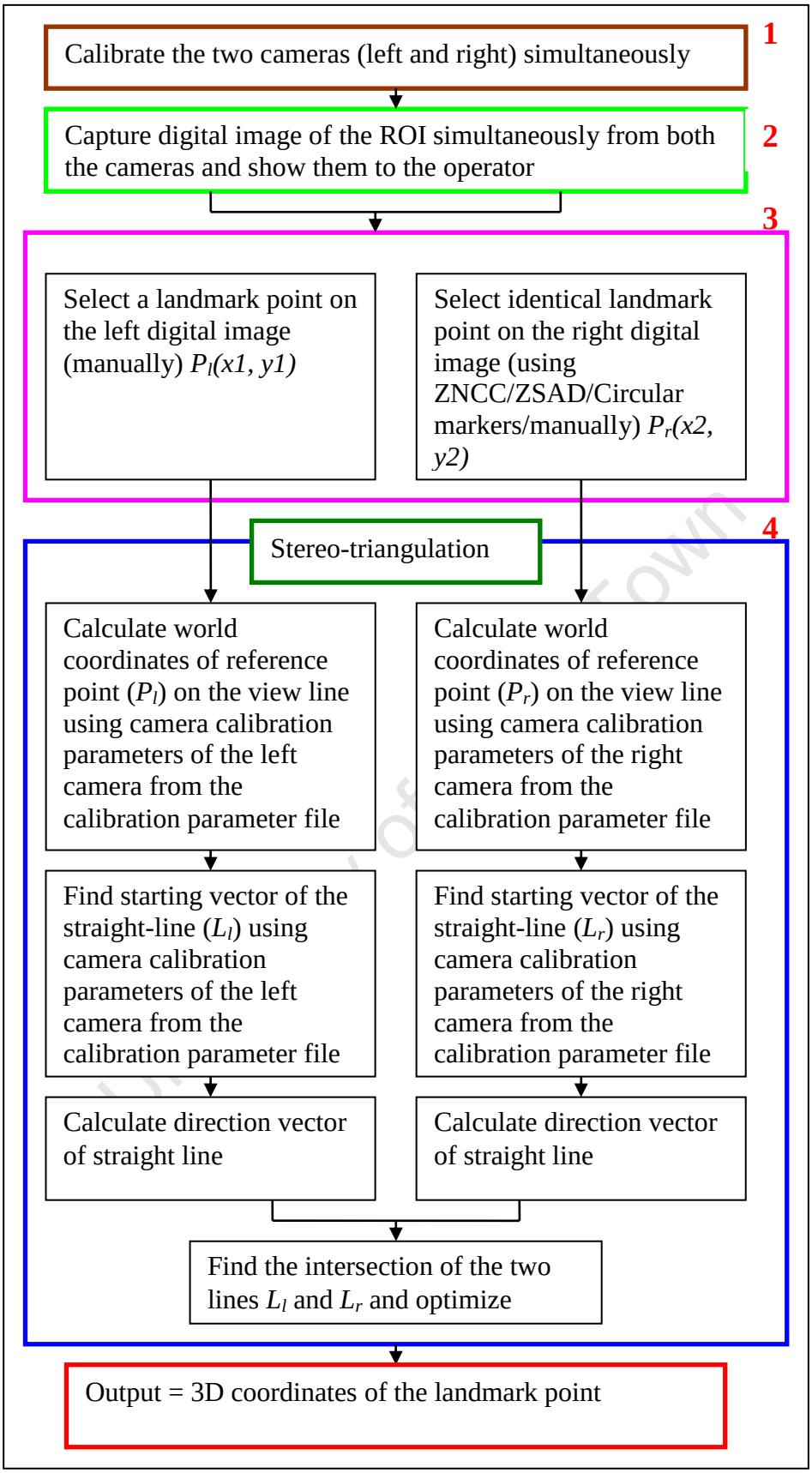

Figure 7.4: Implemented process using IVT functions and applications.

8 USER INTERFACE FOR THE APPLICATION

The user interface and the application were developed using IVT libraries and Visual Studio 2008 Express edition. The application can be used to:

- Calibrate two cameras at a time on either side of the C-arm.
- Simultaneously capture and store digital images from two cameras.
- Determine the length, breadth, and the thickness of the ROI prior to scanning.
- Facilitate landmark selection by using correlation techniques and automatic circular marker detection.
- Determine the distance of the C-arm from the ROI in order to initialize the position of the C-arm at the beginning of the ROI.
- Vary the vertical distance of the object from the cameras manually on the user interface depending upon the size of the patient in order to calculate the disparity internally. The window size or the image patch to be searched, having the landmark point as the centre, on the right digital image can be set in advance.
- The horizontal distance of the C-arm from the cameras can be set in advance.

8.1 Overview of the user interface

The graphical user interface (GUI) is shown in Figure 8.1. The cameras are calibrated using the chessboard pattern by choosing the option “Calibrate Cameras” on the GUI. A particular pair of cameras (on either the left or the right hand side of

the C-arm) can be selected to capture the digital images of the ROI by selecting the option “Capture Images” on the graphical user interface shown in the Figure 8.1.

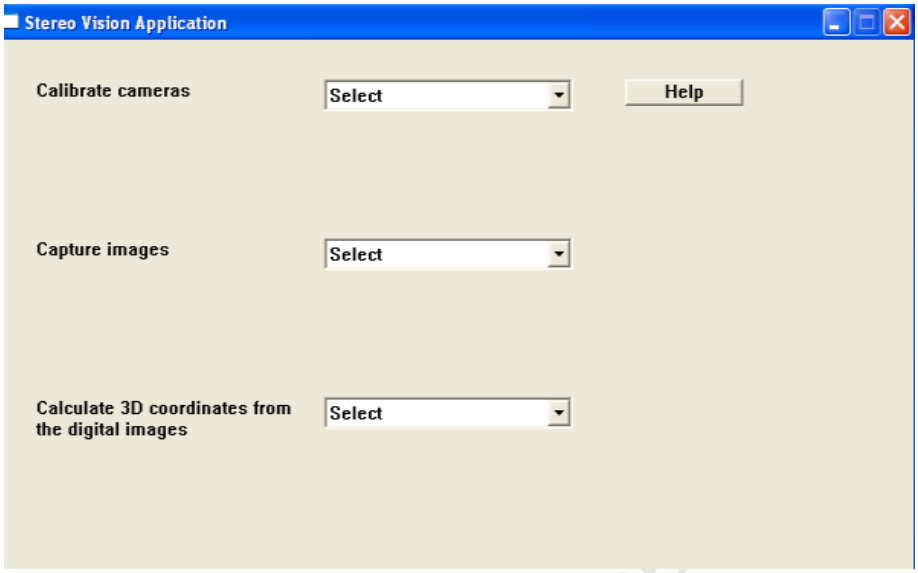

Figure 8.1: Graphical user interface (GUI) of the application for camera calibration, capturing digital images, and for calculating length, breadth, and the thickness of the ROI.

Once the digital images are captured, the landmark points defining the ROI are selected. The operator is assisted in landmark selection with the use of ZNCC, ZSAD, and the automatic circular marker detection (Section 7.6). Any of these methods of point correspondence can be chosen on the user interface by selecting *Method 1* or *Method 2* or *Method 3* from the dropdown as shown in Figure 8.2. *Method 1* corresponds to ZNCC, *Method 2* corresponds to detecting centres of the retro-reflective circular markers, and *Method 3* corresponds to ZSAD.

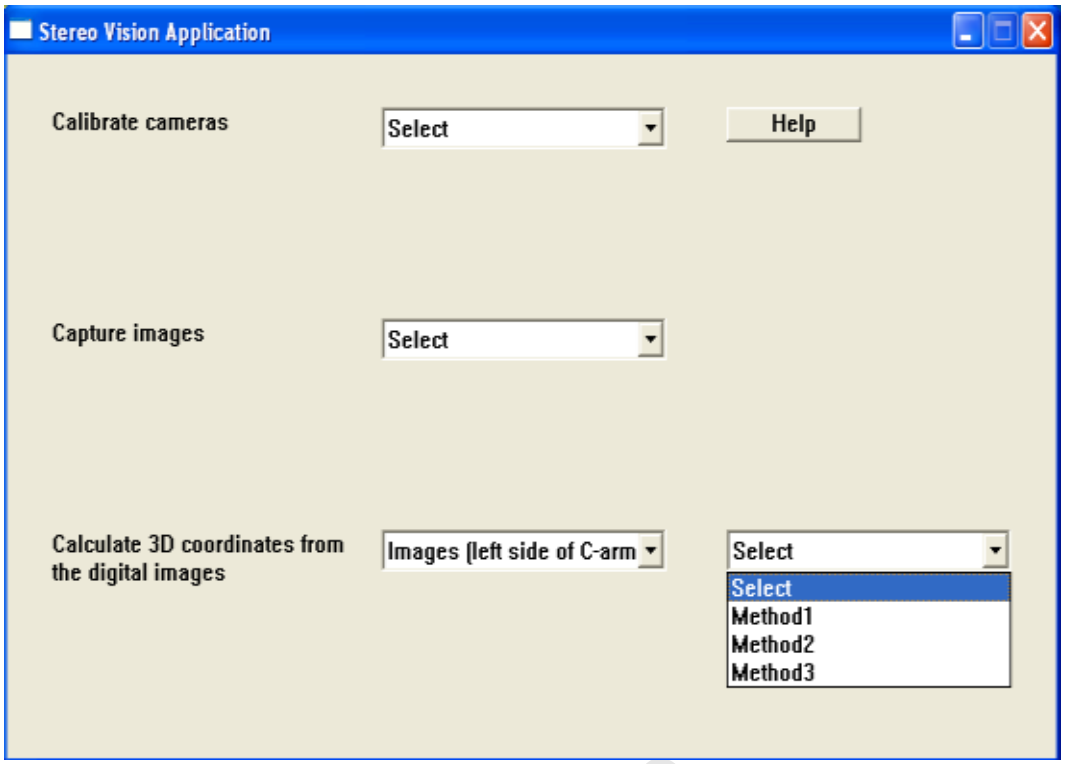

Figure 8.2: GUI of the application. Method 1 corresponds to ZNCC, Method 2 corresponds to detecting centres of the retro-reflective circular markers, and Method 3 corresponds to ZSAD.

8.1.1 ZNCC as matching technique for point correspondence

After selection of *Method 1*, the digital images will appear on the GUI (Figure 8.3). The landmark points defining the ROI are selected one by one on the left digital image. The best match for the landmark point is then searched for along the epipolar line (Section 5.2.4.2) using ZNCC. The operator can manually refine the correlated match by manually clicking the matched landmark point on the right digital image if the best match found with ZNCC fails due to ambiguity, occlusion, or lack of texture.

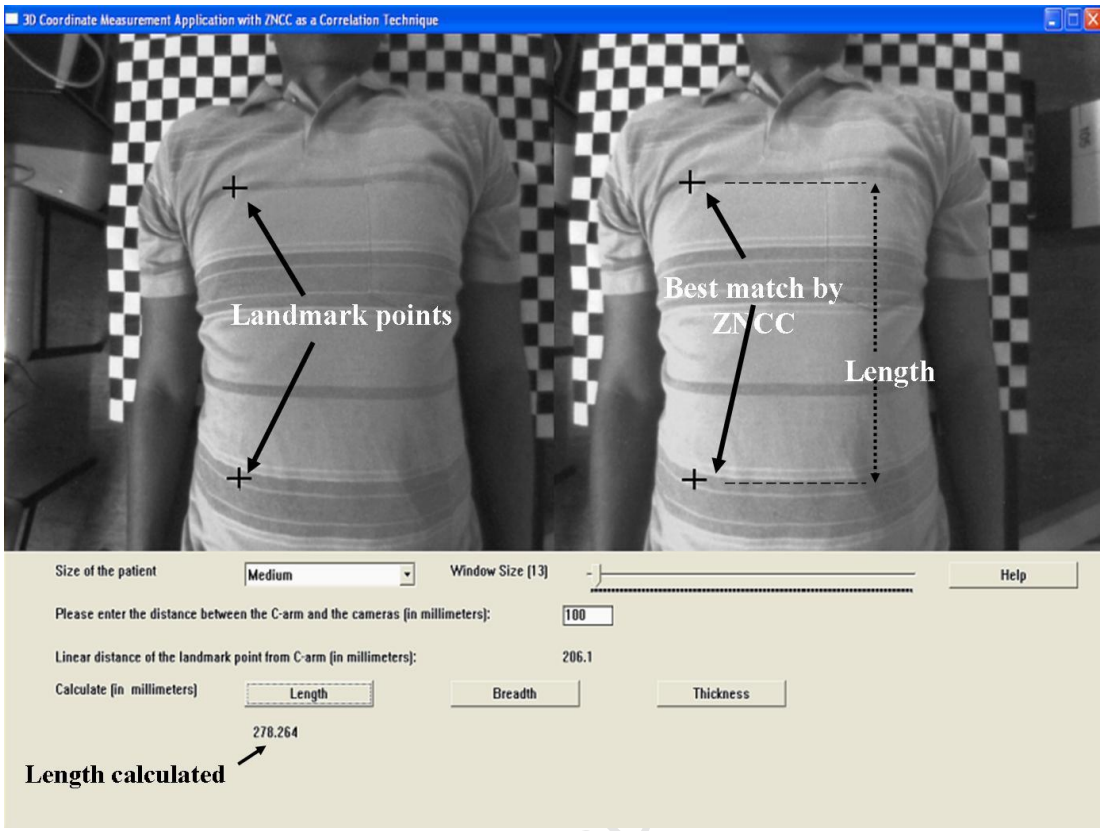

Figure 8.3: GUI for image matching using ZNCC.

8.1.2 ZSAD as matching technique for point correspondence

After selection of *Method 3* from the dropdown box in Figure 8.2, the steps given in Section 8.1.1 for ZNCC are also applicable to ZSAD.

8.1.3 Using retro-reflective markers on the ROI

Four to six retro-reflective markers are placed on the object of interest to indicate the landmark points defining the length, breadth, and the thickness of the ROI. *Method 2* is selected on the GUI. The GUI for automatically detecting the retro-reflective circular markers is shown to the user (Figure 8.4). The markers and their centres are detected automatically using the Hough circle transform (Section 5.3.2). The

operator can manually refine the landmarks or select other landmark points on both the digital images if the circular markers are not clearly visible or are not detected on the digital images. The digital images are not corrected for distortion, as distortion correction was found to impair the detection of circular markers using the Hough circle transform.

Figure 8.4: GUI for detecting centres of the retro-reflective markers using the Hough circular transform. The number of circles, their radius (minimum and maximum), and Canny's low and high threshold values can be changed by the user. The centres of the circular markers are used as landmark points; however, the operator can manually refine the landmark points if required.

Retro-reflective circular markers of any size can be used; however, the markers chosen should be clearly visible in the digital images. The radius (minimum and maximum) of the markers can be set by the user. This feature is provided to increase the utility of the application. The number of circular markers to be used might vary

depending upon the ROI and can be set in the GUI. The default values for Canny's low and high threshold values (Section 5.3.2 and Figure 5.13) gave good results for most of the cases but they can also be changed on the GUI if required.

8.1.4 Determining the ROI

Once the operator manually confirms the selection of the best match on the right digital image for two landmark points clicked on the left, the *length*, *breadth* or *thickness* button is clicked (Figure 8.3 and Figure 8.4), in order to calculate the corresponding value defined by the two landmarks. The value (determined by calculating the distance between the 3D coordinates of the selected landmark points) will appear on the GUI. For determining the thickness of the ROI, one landmark point should be selected on the ROI while another landmark point should be selected on the Statscan patient trolley. A chessboard pattern may be used on the trolley in order to assist landmark selection, as shown in Figure 8.3 and Figure 8.4. The application calculates the thickness by finding the vertical distance between the two landmark points.

The size of the patient can be set in advance. Depending upon patient size (small, medium, large, and extra-large), a rough estimate of the depth of the landmark points from the cameras is calculated. As mentioned in Section 5.2.4.1, if a disparity range is set, the corresponding landmark points will be searched for only in the given disparity range. This reduces the search area for the area based matching techniques. The function available in the IVT image-processing library is used to calculate the

rough disparity range based on the probable depth of the landmark point estimated from the size of the patient (Section 5.2.4.1). A buffer value was determined based on experiments conducted with different pairs of digital images of human subjects to examine the variation of disparity from the rough disparity range. The disparity range (maximum and minimum value) was determined by adding and subtracting the buffer value from the rough disparity.

The window size or image patch in which the search along the epipolar line takes place can also be set in advance. A window size of 13 X 13 (set as default) was used for all the tests conducted in Chapter 9. The landmark point selected on the left digital image is the centre of the image patch to be searched on the right digital image. If the quality of the digital images is not good and the image patch selected for the search does not provide enough information about the landmark point to be searched, area based matching might fail. Window size also plays an important role in searching for the best match on the right digital image for the landmark point selected in the left digital image. The window size chosen should be large enough so that the reference landmark point and the pixels around it have unique and sufficient information for identifying the best match. Large window size might increase the probability of mismatch, however. Therefore, an appropriate window size should be selected for searching area-based matching (Roma *et al*., 2002).

The horizontal distance of the cameras from the C-arm is measured manually and is updated in the textbox provided on GUI as shown in Figure 8.3 and Figure 8.4. The distance of the C-arm from ROI is equal to the sum of the distance of the cameras

from the ROI (calculated using the algorithm) and the distance of the cameras from the C-arm (entered manually by the operator on GUI). This distance is also displayed on the GUI.

9 EXPERIMENTS AND RESULTS

The application developed was tested on test objects as well as on different human subjects. The first set of experiments compared distance between landmarks, obtained manually and using stereo-photogrammetry, while the second compared image/pixel coordinates and 3D coordinates in mm. The two approaches were taken, because the first allowed a physical "ground truth" measurement to indicate the accuracy of stereo-photogrammetry, while the second indicated the accuracy of the matching methods when used in conjunction with stereo-photogrammetry.

9.1 3D distance measurement on test objects

The cameras were calibrated using a planar chessboard pattern shown in Figure 7.3. The camera pair captured digital images of different test objects shown in Figure 9.1. The 3D coordinates of the landmark points were determined using the application.

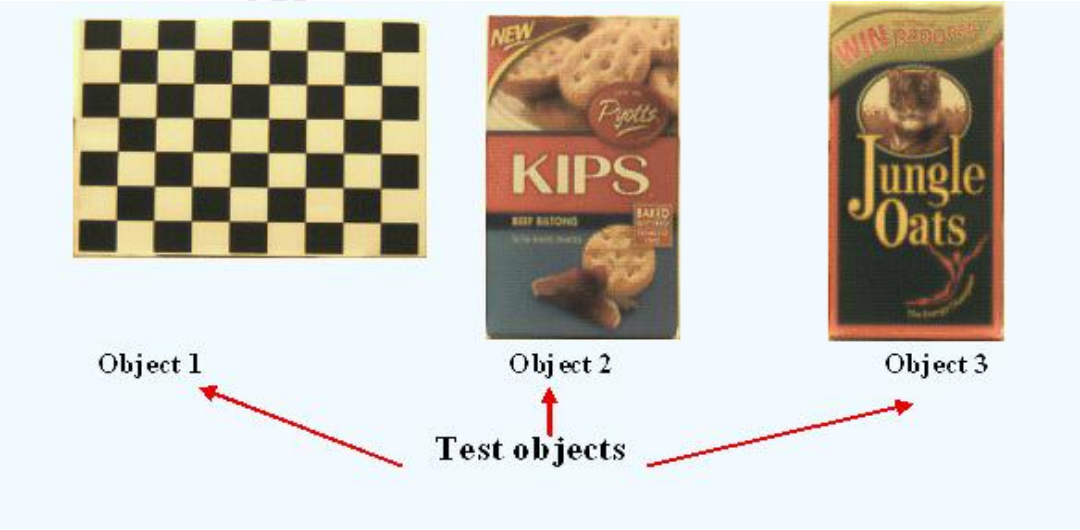

Figure 9.1: Objects used for testing the 3D coordinate measurement application.

In precision tests for accuracy, direct measurements on test objects of interest using hand held rulers, callipers, or soft measuring tapes to obtain linear distances have been used extensively for assessing the reliability of indirect methods (stereo-photogrammetry in this case) of obtaining these distances (Aung *et al.*, 1995; Grobbelaar, 2005; Mutsvangwa, 2009; Weinberg *et al.*, 2006). The accuracy of 3D coordinates on the test objects was determined by calculating the distance between two landmark points and comparing it to the actual distance measured using a hand held ruler. A hand held ruler was used in this case, as the distances were small and the landmark points were quite distinct on test objects. The landmark points were selected based on ease of identification using manual selection and stereo-photogrammetry.

The distances between landmark points selected on the test objects were in the range of 30 to 200 millimetres. The number of distances measured varied from two to thirteen depending upon the landmark points clearly visible on the test objects. Thirty-six distinct distances were measured in total for all the test objects and each distance was measured three times for precision and averaged. The mean absolute difference between actual distances (measured manually) and observed distances (using the application), standard deviation, and mean difference as a percentage of actual distance were computed for all the distances (108 distances as each of the thirty-six distances were measured three times) on all the test objects. The results are shown in Table 9.1.

Table 9.1: Distances between landmark points on three different test objects. The distances between the landmark points varied from 30 to 200 millimetres. Thirty-six distinct distances (two to thirteen distances on each test object) were measured on each object, and each distance was measured three times.

Mean absolute difference between actual and observed distances; range in brackets	Standard deviation (mm)	Mean difference as a percentage of actual distance; range in brackets
0.7 (0 to 3 mm)	0.5	1.7 (0.01% to 9.7 %)

9.2 3D coordinate measurement application on human subjects

Experiments were conducted to verify the accuracy of the length, breadth, and thickness of the ROI obtained using the developed application on human subjects. The results were verified by comparing them with the distances measured manually using calipers (Aung *et al*., 1995; Grobbelaar, 2005; Mutsvangwa, 2009; Weinberg *et al*., 2006). Initially the experiments were conducted on a subject by manually selecting corresponding landmark points on both the digital images. In the later experiments, matching techniques (Section 8.1.1, 8.1.2, and Section 8.1.3) were used on five subjects to find the best match.

9.2.1 Manual selection of landmark points

Corresponding landmark points were selected manually on both the digital images. The length, breadth, and the thickness of the ROI were compared with the actual distances measured manually using calipers. A comparison of inter-landmark distances for a human subject is given in Table 9.2, for manually selecting the corresponding points. The length, breadth, and thicknesses compared were in the

range of 20 to 240 millimeters. Each landmark was selected and each distance was measured three times.

Table 9.2: Distances between landmark points on a human subject. Each distance (length, breadth, thickness) was measured three times during testing, i.e. nine distances were measured .

ROI	Mean absolute difference between actual and observed distances ; range in brackets	Standard deviation (mm)	Mean difference as a percentage of actual distance ; range in brackets
Length	2.2 (2 to 2.6 mm)	0.3	1.1 (0.9 to 1.2 %)
Breadth	7.9 (6.4 to 9.1 mm)	1.4	4.0 (3.2 to 4.5 %)
Thickness	5.3 (0.6 to 7.6 mm)	4.0	2.3 (0.2 to 3.2 %)

9.2.2 Accuracy of the best match found using ZNCC and ZSAD matching techniques

A landmark point was selected in the left digital image. The best match for the same landmark point was searched for in the right digital image using ZNCC and ZSAD respectively. The match was refined manually on the right digital image, for five image sets. Approximately 16 landmark points from each of the set of digital images of a subject (total 82 for sets of digital images from 5 subjects) were selected on the digital image on the left hand side. The best match for the same landmark point was found using ZNCC and ZSAD methods on the right digital image. The matching points may also be chosen manually and the user may override the matching points obtained using the aforementioned ZNCC and ZSAD methods.

In the experiment conducted, the manually refined landmark point (after the landmark points were searched for using ZNCC and ZSAD methods) on the right

digital image for the landmark point clicked on the left digital image was considered close to the true match. The image coordinates (in pixels) of the landmark points found in the right image (after manual selection in the left image); using ZNCC and ZSAD (uncorrected) were compared with the image coordinates of the manually refined match on the right digital image. The 3D coordinates (in mm) of the landmark points obtained using ZNCC and ZSAD were also compared with the 3D coordinates of the manually overridden best match in the right digital image. The mean differences, standard deviation in image coordinates (in pixels), and 3D coordinates (in mm) were as shown in Table 9.3 and Table 9.4.

Table 9.3: Differences between the best match found using ZNCC, and manually marked landmark points. Sixteen landmark points were selected on each of the five different subjects.

	Pixel difference (manual and ZNCC)		**Difference in mm (manual and ZNCC)**		
	X (pixels)	**Y (pixels)**	**X(mm)**	**Y(mm)**	**Z(mm)**
Range of absolute differences	0 to 2.7	0 to 2.1	0 to 6.3	0 to 3.5	0 to 19.9
Mean error	0.5	0.6	0.8	0.9	4.3
Standard deviation	0.4	0.4	1.0	0.7	3.7

Table 9.4: Differences between the best match found using ZSAD, and manually marked landmark points. Sixteen landmark points were selected on each of the five different subjects.

	Pixel difference (manual and ZSAD)		**Difference in mm (manual and ZSAD)**		
	X(pixels)	**Y(pixels)**	**X(mm)**	**Y(mm)**	**Z(mm)**
Range of absolute differences	0 to 10.8	0 to 2.1	0 to 6.8	0 to 5.1	0 to 22.8
Mean error	0.6	0.6	0.8	0.9	4.2
Standard deviation	1.2	0.4	1.1	1.0	4.8

Bland and Altman (1986) suggested a graphical method for evaluating whether the measurements obtained using two different methods are in agreement especially when the true values of comparison are not available. In this method, a graph should be plotted using values of differences (Y-axis) against the mean values (X-axis). Accordingly, graphs were plotted using 3D coordinates in X,Y and Z directions, values of differences (Y-axis) against the mean values (X-axis) for the assessment of agreement between the matching methods (ZNCC and ZSAD) and the corresponding values obtained by manually refining the results of these matching methods. The graphs are shown in Figure 9.2, Figure 9.3, Figure 9.4, Figure 9.5, Figure 9.6, and Figure 9.7 respectively. The limits within which 95% of measurements lie in the X, Y and Z directions using manually refined and correlation methods from the graphs in Figure 9.2, Figure 9.3, Figure 9.4, Figure 9.5, Figure 9.6 and Figure 9.7 are summarized in Table 9.5. The sign of the difference was preserved during calculations for determining the maximum and minimum limits of agreement.

Table 9.5: The limits within which 95% of measurement differences lie in the X, Y and Z directions using matching methods and manual refinement. The maximum limit corresponds to Mean + 1.96 Standard Deviation and the minimum limit corresponds to Mean - 1.96 Standard Deviation.

	Upper limit (Manually refined and ZNCC) (mm)	**Lower limit (Manually refined and ZNCC) (mm)**	**Upper limit (Manually refined and ZSAD) (mm)**	**Lower limit (Manually refined and ZSAD) (mm)**
X direction	2.7	-2.5	3	-2.5
Y direction	2.4	-2.3	2.4	-3.0
Z direction	11.3	-11.0	12.0	-13.2

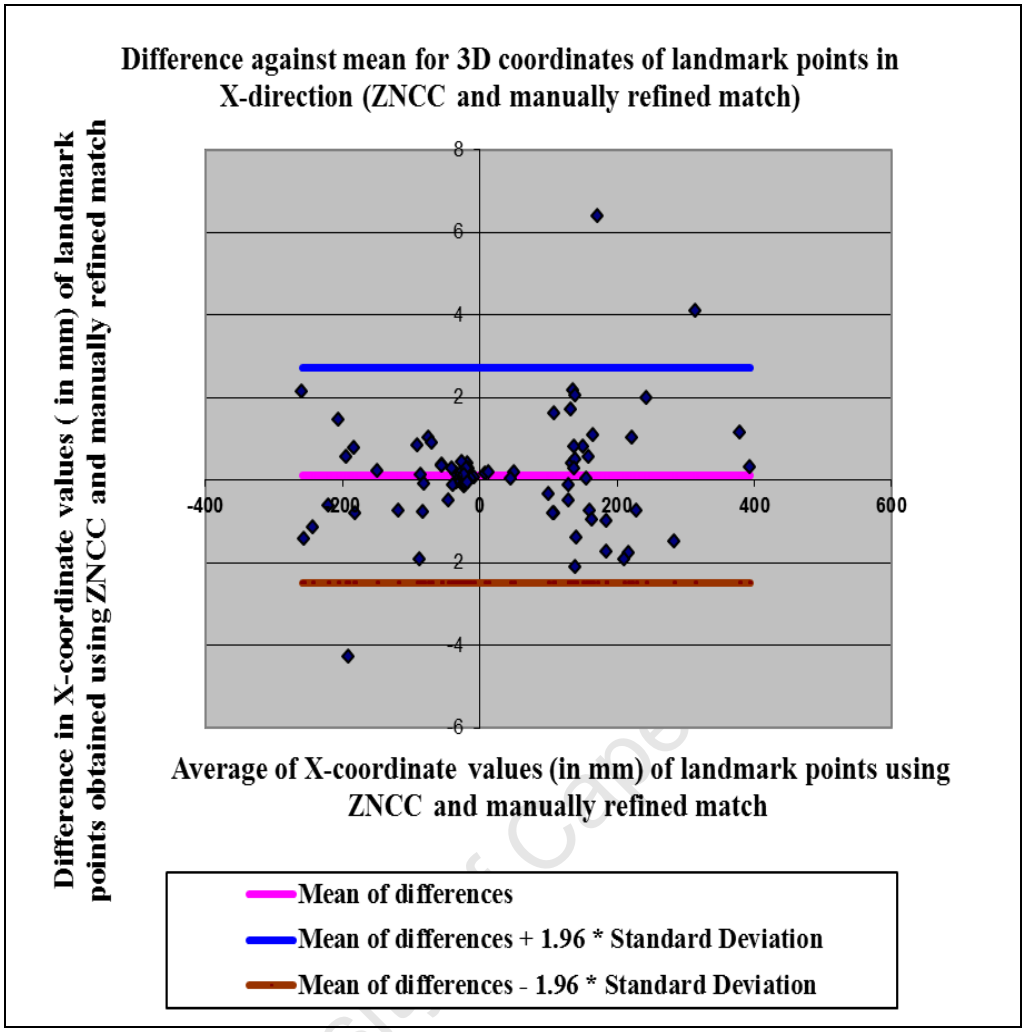

Figure 9.2: Difference against mean plot in X-direction for 3D coordinates obtained using ZNCC and manually refined match.

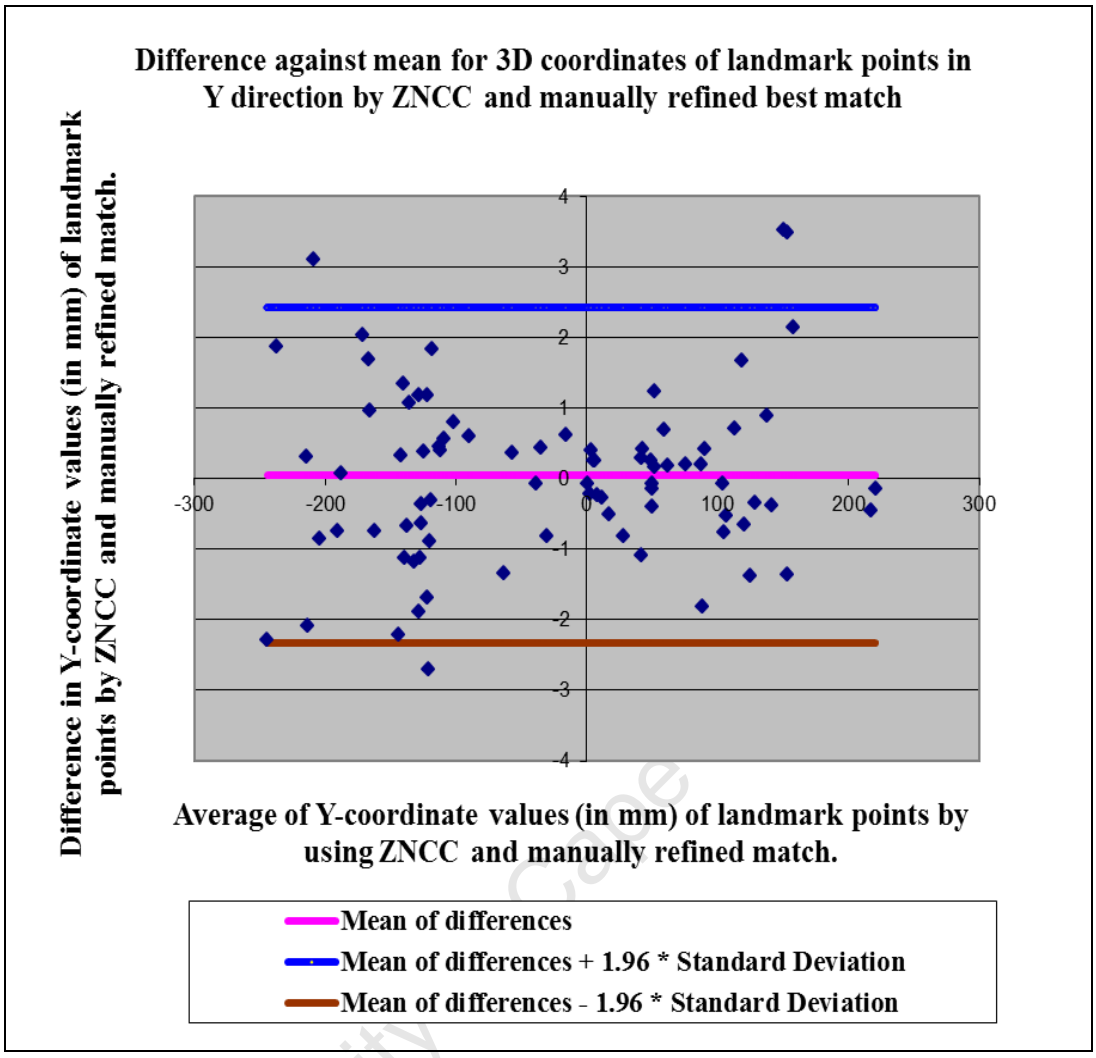

Figure 9.3: Difference against mean plot in Y-direction for 3D coordinates obtained using ZNCC and manually refined match.

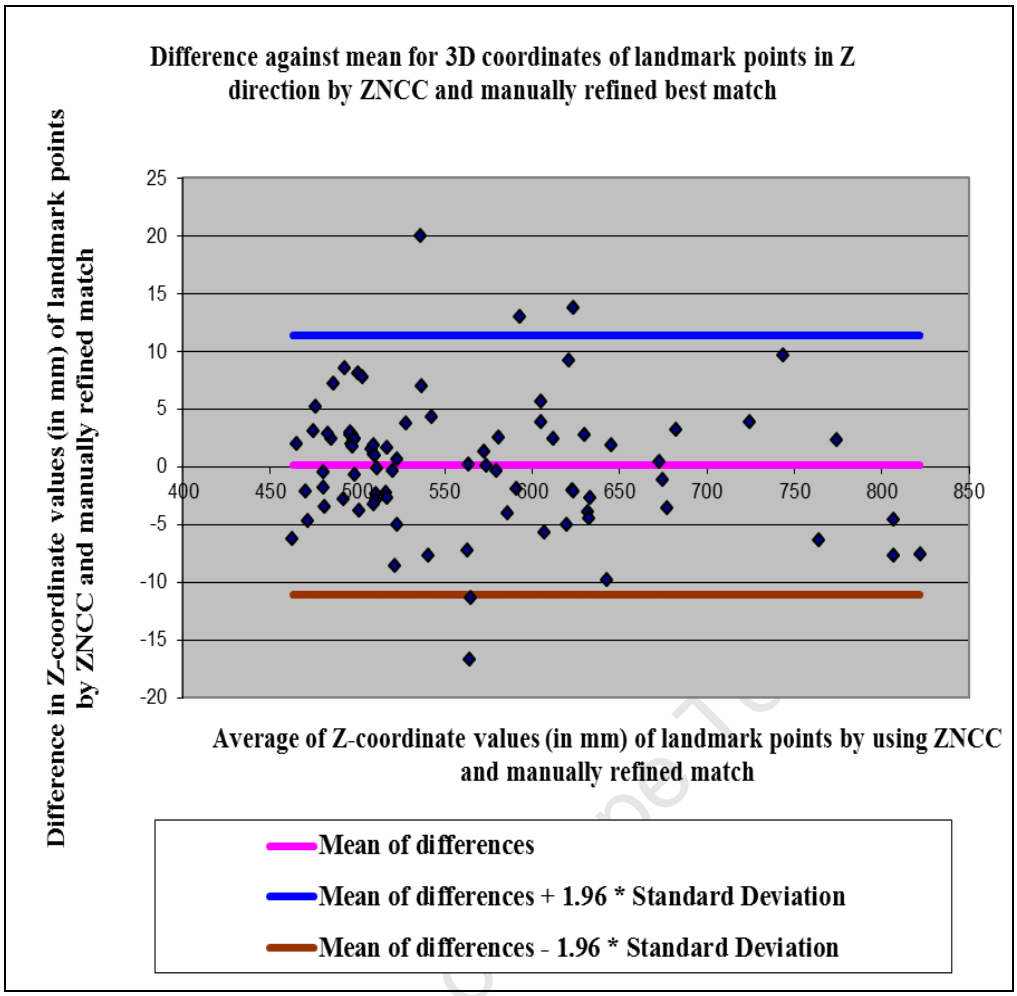

Figure 9.4: Difference against mean plot in Z-direction for 3D coordinates obtained using ZNCC and manually refined match.

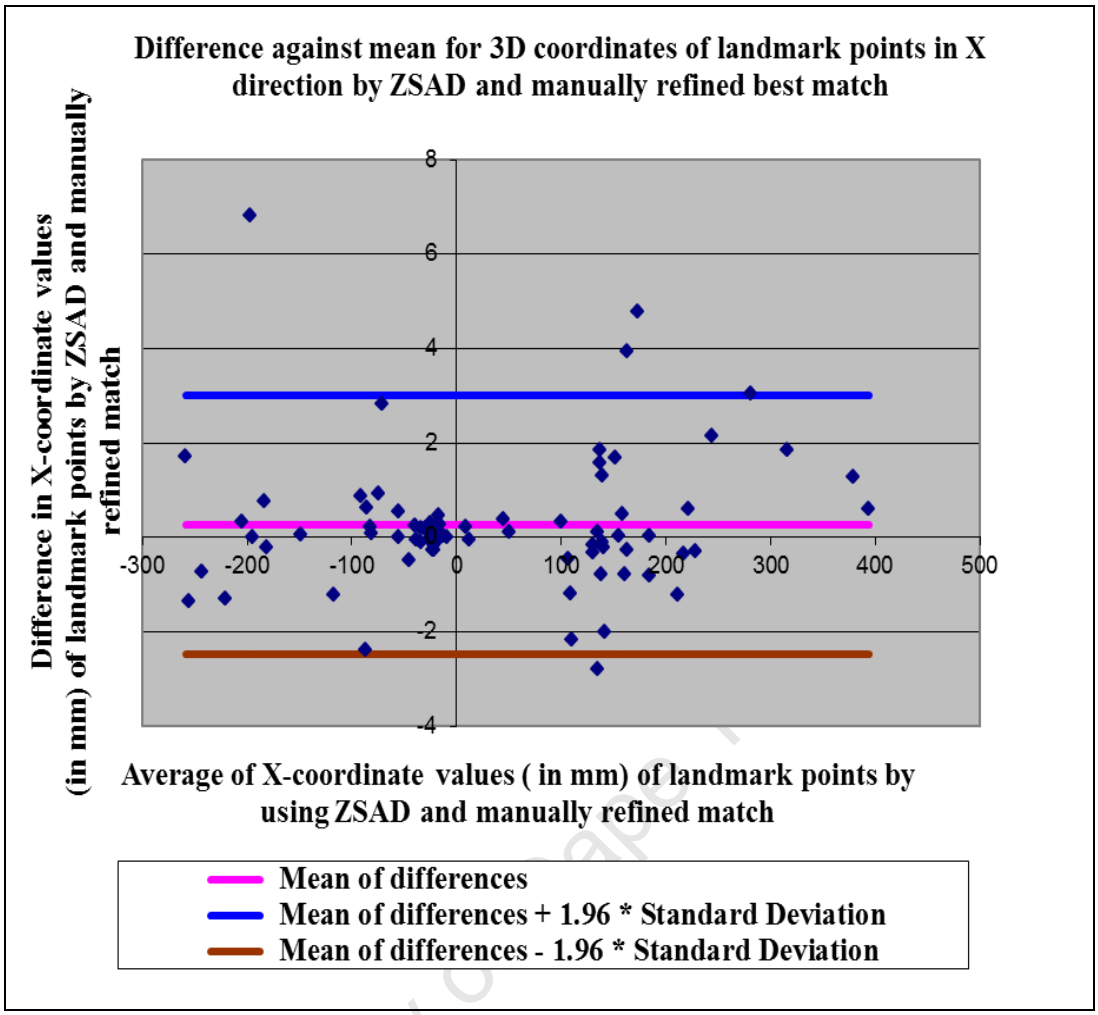

Figure 9.5: Difference against mean plot in X-direction for 3D coordinates obtained using ZSAD and manually refined match.

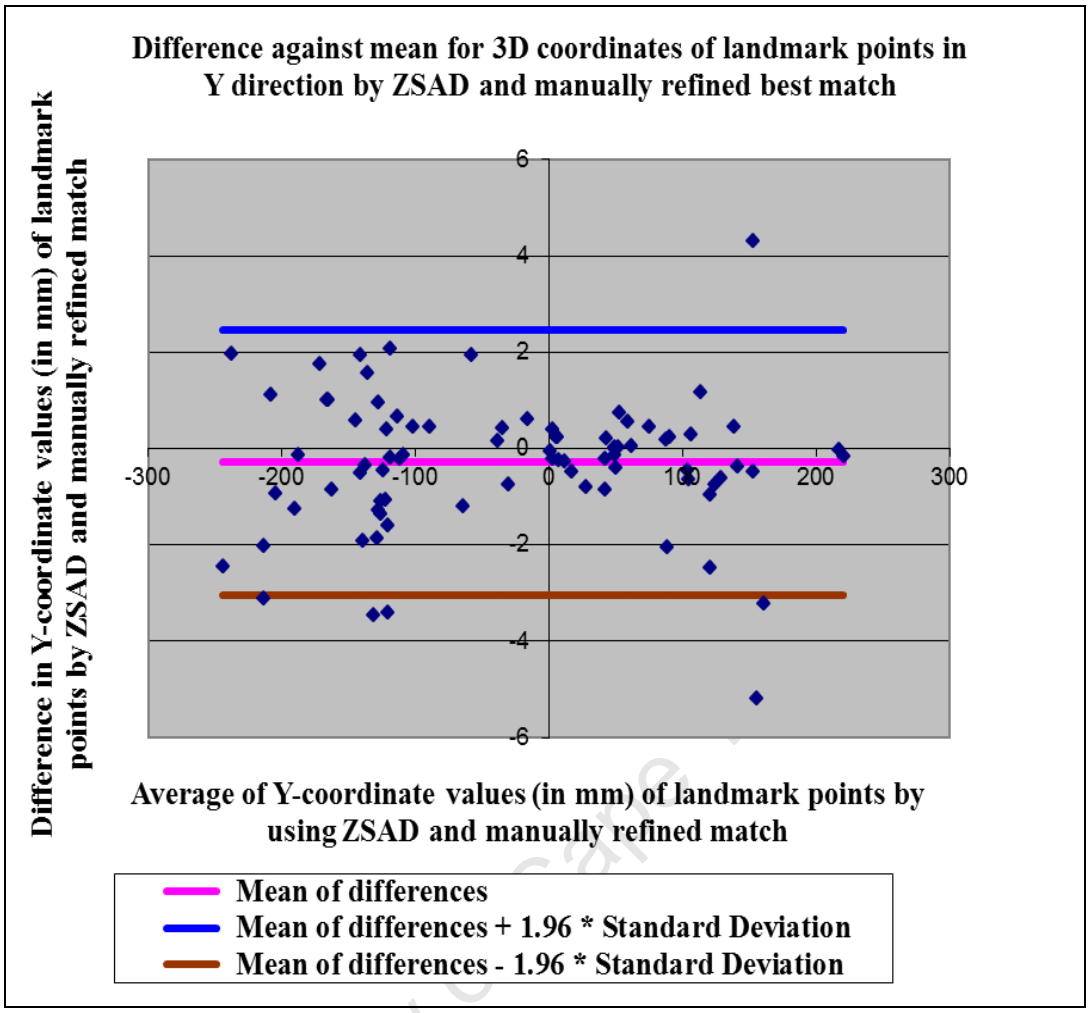

Figure 9.6: Difference against mean plot in Y-direction for 3D coordinates obtained using ZSAD and manually refined match.

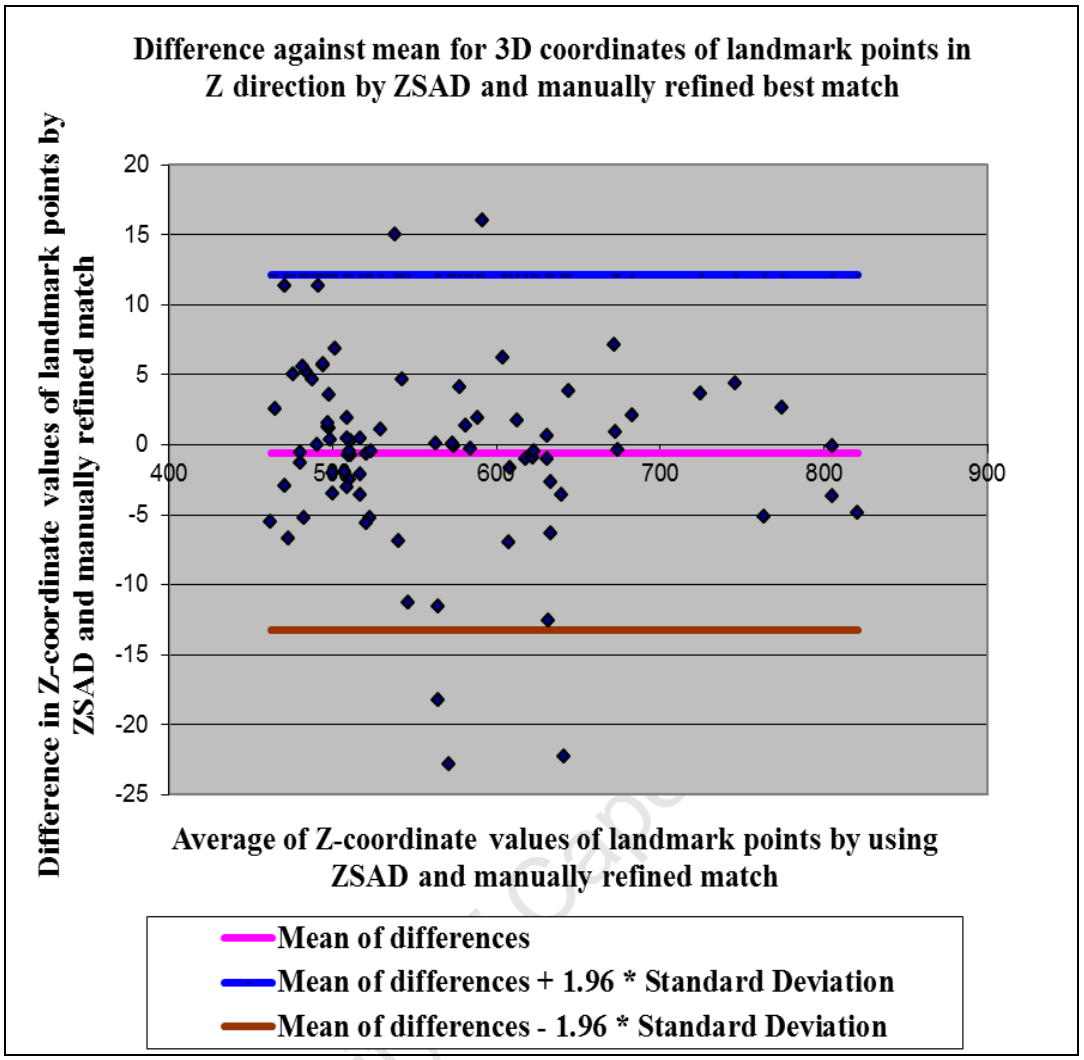

Figure 9.7: Difference against mean plot in Z-direction for 3D coordinates obtained using ZSAD and manually refined match.

9.2.3 Accuracy of automatic circular landmark detection

The centres of the retro reflective markers (four circular markers of 32 mm diameter) were detected automatically (using the Hough circle transform) on digital images (both left and right) of five different subjects. The matching points (centres of the markers in this case) may also be chosen manually and the user may override the matched points obtained using the aforementioned method.

The accuracy of centres of the circular markers obtained using the Hough circle transform was determined by comparing the pixel coordinates and the 3D coordinates in mm of the automatically detected centres with centres refined manually (after automatic selection using the Hough transform). The results are shown in Table 9.6 below.

Table 9.6: Circular marker detection on five different image pairs. Four circular markers were used on each subject.

	Image coordinates of the centre of the retro-reflective circular markers (manually refined and Hough circle transform)		**Difference in mm in 3D coordinates (manually refined and Hough circle transform)**		
	X (pixels)	**Y(pixels)**	**X(mm)**	**Y(mm)**	**Z(mm)**
Range of absolute differences	0 to 5	0 to 6	0 to 3.4	0 to 5.8	0 to 10.4
Mean absolute difference	1.1	1.5	1.2	1.5	4.2
Standard deviation	1.2	1.5	1.0	1.6	3.8

In addition, Bland and Altman graphs were plotted using 3D coordinates in X,Y and Z directions, values of differences (Y-axis) against the mean values (X-axis) for the assessment of agreement between the centres of the circular markers by Hough circle transform and the corresponding values obtained by manually refining the results of the Hough circle transform. The graphs are shown in Figure 9.8, Figure 9.9, and Figure 9.10.

The limits within which 95% of measurement differences between the Hough circle transform only and manual refinement of the Hough transform results lie are

summarized in Table 9.7. The sign of the difference was preserved during calculations for determining the limits of agreement.

Table 9.7: The limits within which 95% of measurement differences lie in the X, Y and Z directions using centres of the circular markers selected by the Hough circle transform and refined manually.

	Upper limit (Manually refined and Hough circle transform) (mm) (Mean + 1.96 Standard Deviation)	**Lower limit (Manually refined and Hough circle transform) (mm) (Mean - 1.96 Standard Deviation)**
X direction	2.8	-3.4
Y direction	4.2	-4.6
Z direction	10.9	-11.6

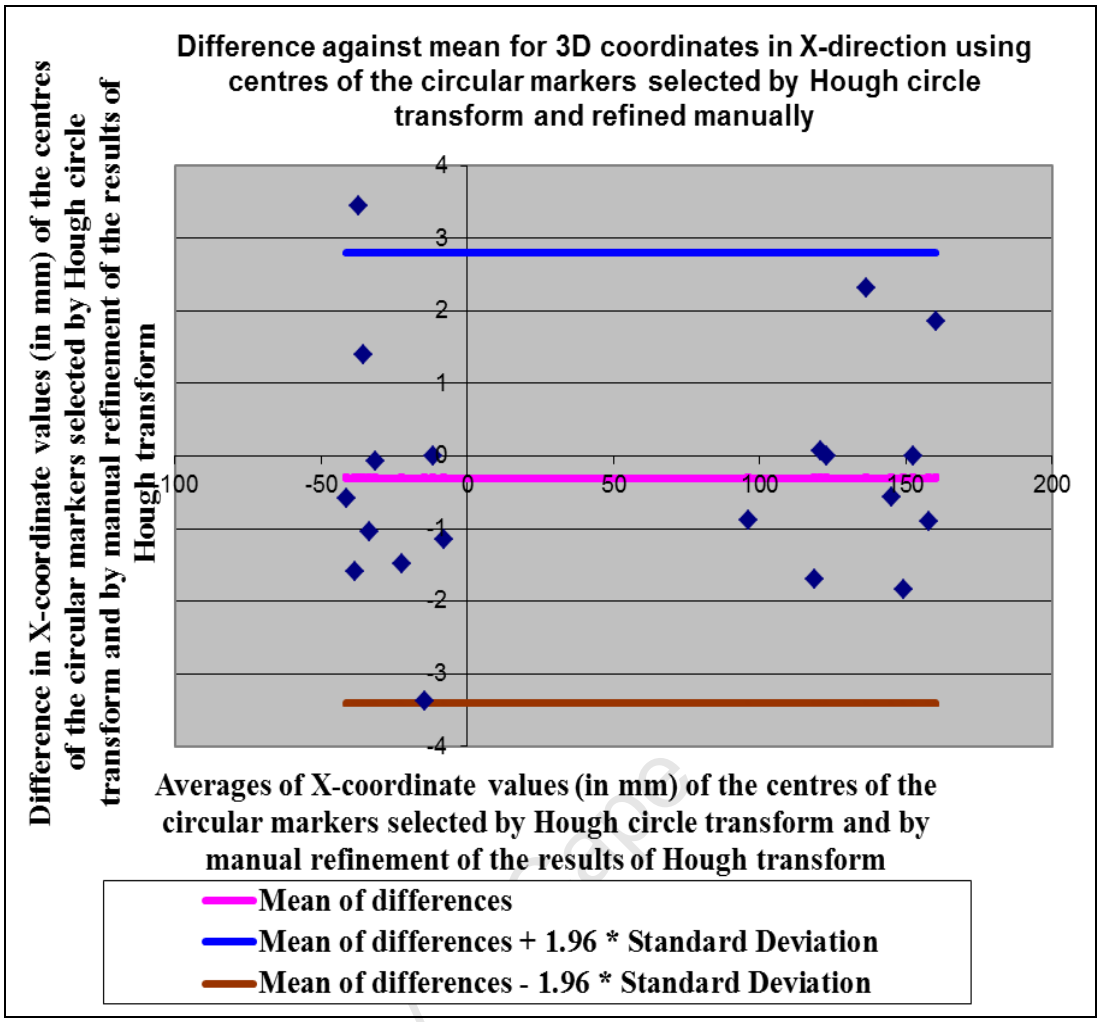

Figure 9.8: Difference against mean plot in X-direction for 3D coordinates of the centres of circular markers selected by Hough circle transform and refined manually.

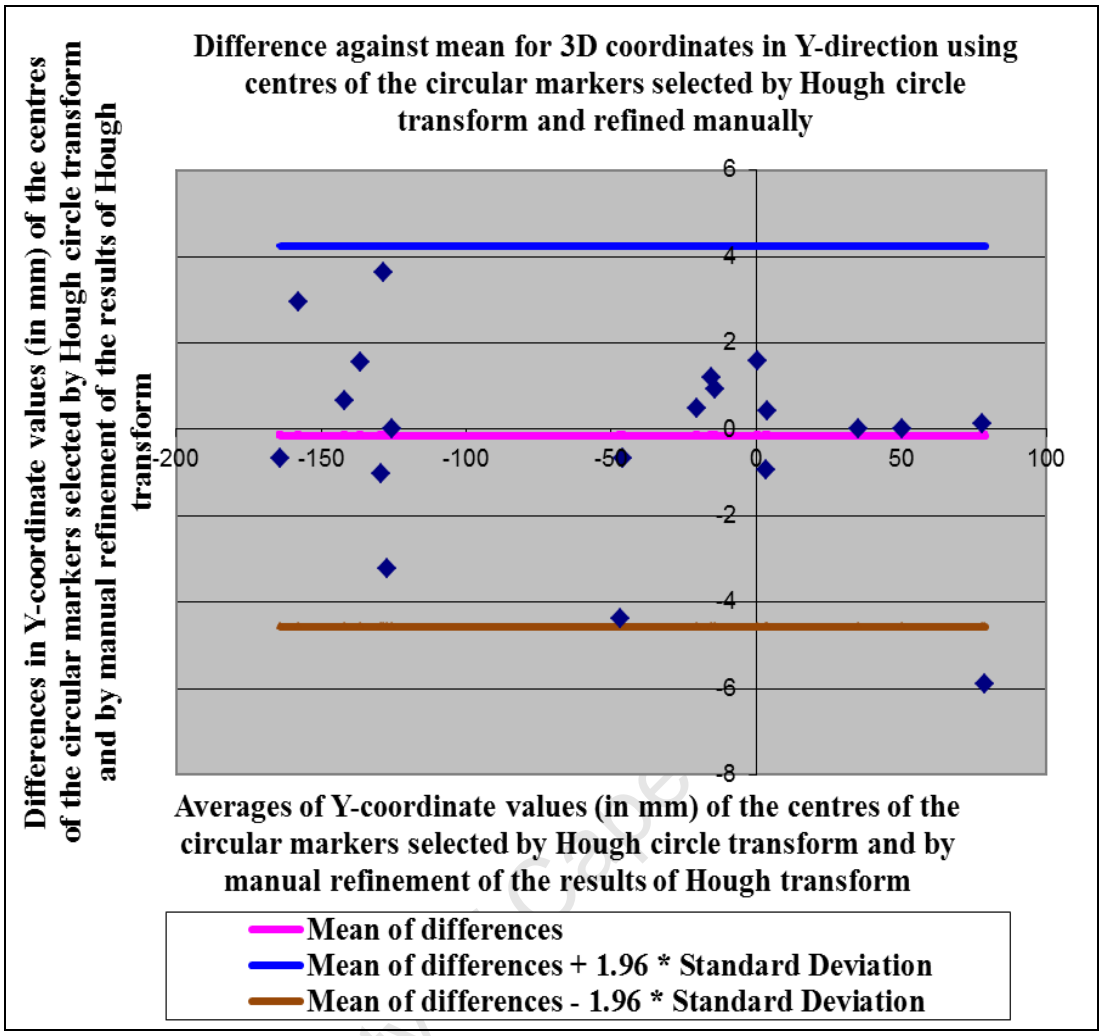

Figure 9.9: Difference against mean plot in Y-direction for 3D coordinates of the centres of circular markers selected by Hough circle transform and refined manually.

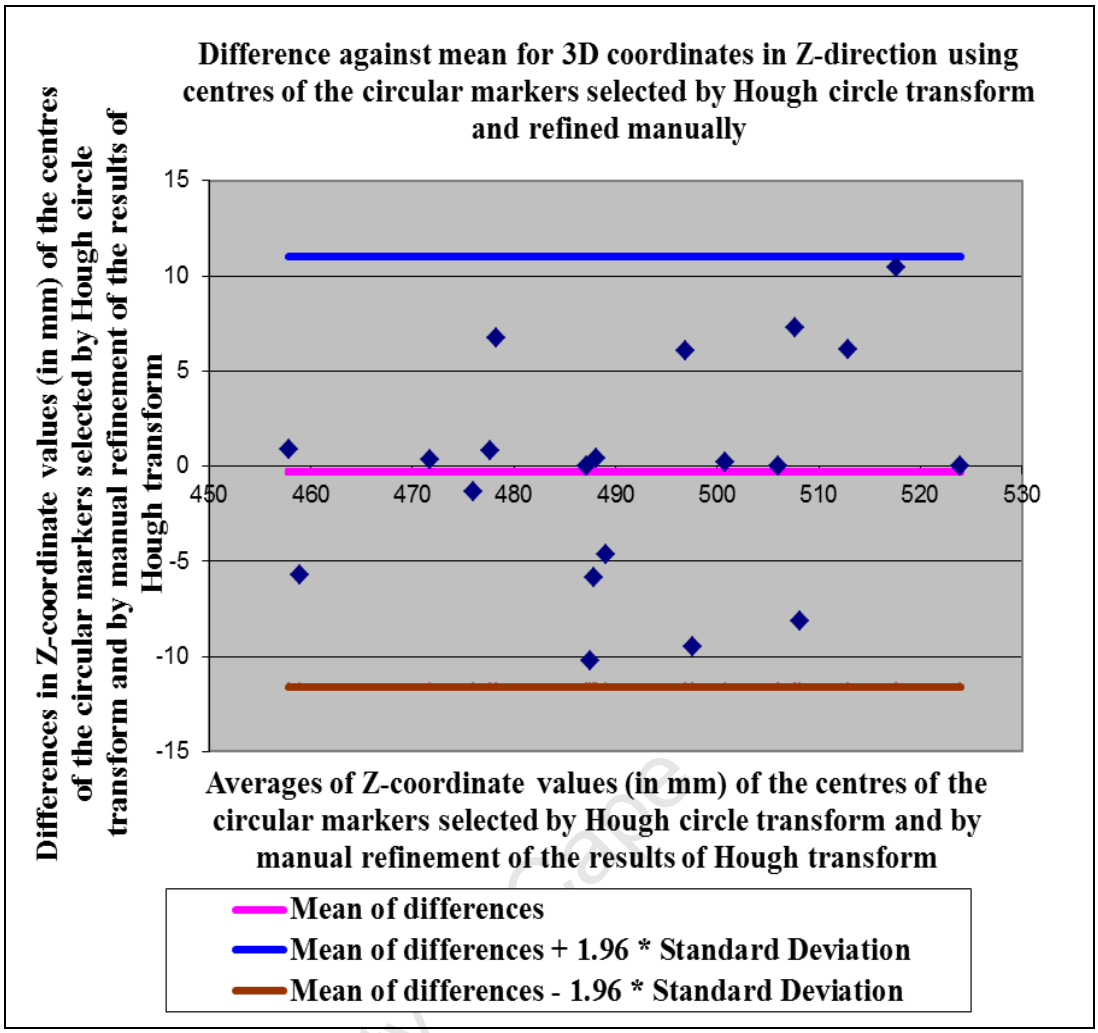

Figure 9.10: Difference against mean plot in Z-direction for 3D coordinates of the centres of circular markers selected by Hough circle transform and refined manually.

10 DISCUSSION

Methods were developed to locate regions of interest on the body prior to X-ray imaging. Stereo-photogrammetry and landmark matching techniques were successfully employed to achieve this goal. The methods were tested using objects as well as human subjects. Image processing software was developed to perform matching using three techniques (ZNCC, ZSAD, and circular markers) and to calculate length, breadth thickness of the ROI from a stereo pair of images.

10.1 3D measurement on test objects and on human subjects without using matching techniques

Testing manual selection of landmarks on both stereo images revealed that ROIs were located well within the permissible error of less than 13 mm, for both inanimate objects and human subjects. The proposed stereo-photogrammetry implementation is therefore an acceptable method to determine the 3D coordinates of landmarks defining the ROI.

From the statistical comparison, the mean absolute difference between actual and observed distances were less for test objects (0.7 mm) than for human subjects (in the range of 0.6 mm to 8 mm) (Table 9.1 and Table 9.2 respectively).

The distances were measured using callipers on the human subject and using a ruler on the inanimate objects. The callipers were in direct contact with the subject's skin

or clothing, deforming the shape and inducing possible errors in direct measurements. Landmark identification was easy on test objects on which many distinct landmark points were available. However, it is difficult to manually identify and select identical landmark points on digital images of human subjects where the area of interest is smooth, without texture, and lacking in distinct features. In addition, movement of the subject between direct measurement and measurement by stereo-photogrammetry would induce errors.

10.2 Stereo matching techniques

The selection of corresponding landmarks in stereo-pairs of digital images was facilitated by implementing ZNCC and ZSAD, and using circular markers on the subjects prior to capturing their digital images.

10.2.1 ZNCC and ZSAD matching techniques

Both ZNCC and ZSAD techniques gave mean absolute pixel errors of approximately one pixel in the X and Y directions when compared to the manually refined match (Table 9.3 and Table 9.4). These results were better than those obtained by Grobbelaar (2005) using cross correlation to match landmarks in a second image to those selected manually in a first image; mean absolute differences in the range of two to six pixels in the X and Y directions were found for facial landmarks.

The 3D coordinates obtained using stereo-photogrammetry by selecting identical landmark points using ZNCC and ZSAD were also compared in the X, Y and Z

directions. The true values of the 3D coordinates of landmark points were not known and it was assumed that the 3D coordinates of the manually refined match represented the actual 3D coordinates. The mean difference between the manual method and the matching methods (ZNCC and ZSAD) were within the permissible limit of 13 mm (Table 9.3 and Table 9.4). Others (Banks and Corke, 2001; Sun and Peleg, 2004) have shown that these methods are comparable in accuracy. A direct comparison of the results could not be done with the results obtained by Sun and Peleg (2004) and Banks and Corke (2001) as they employed different techniques (such as analysing depth map images) to assess the precision of these area based matching techniques.

The upper and lower limits defining the 95% confidence interval for measurement differences in the X, Y, and Z directions for 3D coordinates obtained by manual and matching methods were within the permissible limit of 13 mm. The matching techniques are therefore suitable for the proposed solution.

The errors were higher in the Z direction than in the X and Y directions. Other studies have shown that the Z or depth axis is more prone to error (Blostein and Huang, 1987; Kyto *et al.*, 2011; Grobbelaar and Douglas, 2007; Mcvey and Lee, 1982; Meintjes *et al.*, 2002; Mutsvangwa *et al.*, 2009).

Douglas et al (2003), Grobbelaar (2005), and Mutsvangwa et al (2009) suggested that the errors in the Z direction in their research were due to the selection of landmark points at different depths in multiple images. Even small errors in

landmark selection can cause large errors in the X, Y, and Z coordinate values (Azad *et al*., 2008).

Depth accuracy can also be affected by poor camera calibration, which might result if the focal length, distance between the cameras or alignment changes after initial calibration (Blostein and Huang, 1987; Kyto *et al*., 2011; Mcvey and Lee., 1982).

10.2.2 Circular markers as a unique feature for finding corresponding landmark points

Mean absolute differences of less than two pixels were obtained between the automatically detected and manually refined centres in the X and Y direction for circular marker detection (Table 9.6).

The true values of the 3D coordinates of the centre of the circular markers were not known and it was assumed that the 3D coordinates of the manually refined centre of the circular markers represented the actual 3D coordinates. The mean absolute differences between the manual method and the automatic method of detecting circular markers using the Hough circle transform were less than 2 mm in the X and Y directions and less than 5mm in the Z direction (Table 9.6).

The upper and lower limits defining the 95% confidence interval for measurement differences in 3D coordinates of the centres of the circular markers in the X, Y, and Z directions by manual selection and by Hough circles transform were within the permissible limit of 13 mm (Table 9.7). The proposed matching technique (Hough

circle transform for automatically detecting centres of circular markers) is therefore suitable for the proposed solution.

10.3 Processing time

Processing time has not been measured, since the developed application has not been incorporated into the Statscan.

10.4 Cost and ease of use

There are several companies that produce and sell stereo-photogrammetric systems for clinical photographic applications. Dimensional imaging Ltd (Dimensional Imaging Ltd) offers range of 3D products based on passive stereo-photogrammetry for volume difference measurement, surface area measurement, point based measurements (such as distance ratio and angle) or for symmetrical analysis of the surfaces captured by 3D captured systems. 3dMD Inc (3dMd Inc) develops non-invasive 3D surface imaging stereo-photogrammetric tools (active and passive) for creating and accessing digital models of the patient's surface. Precision 3D Limited (Precision 3D Limited) uses stereo-photogrammetry in the area of orthopaedic and custom shoemaking, orthotics, and wound measurement.

The cost of the application thus developed for the research project might be low when compared with the clinical products developed by the companies outlined above. The software product developed requires simple apparatus such as a camera-mounting unit, a chessboard calibration frame, and two or four low cost cameras.

Visual Studio 2008 Express edition is freely available. The IVT library is also freely available under the GNU license.

Camera calibration requires moving the calibration frame in different directions in front of the cameras. It takes less time and less human intervention when compared with camera calibration using a 3D calibration frame. The image capture, calibration, 3D coordinate, and distance measurement techniques are embedded in a single software application to reduce processing time and improve ease of use.

Operator dependency is reduced through the use of matching techniques such as ZNCC, ZSAD, and circular marker detection. Cameras are also calibrated dynamically and the need for landmark point selection is reduced by using the Hough transform to detect the lines automatically in the calibration object.

10.5 Significance of results for Lodox Statscan and Versascan

Experiments conducted showed that stereo-photogrammetry is an acceptable technique for locating ROIs prior to X-ray scanning. Future research needs to be done on ways to use the thickness of the ROI to set the kV and mA levels for obtaining the image.

The proposed solution can also be extended to the Versascan machine to scan smaller ROIs. The camera-mounting unit designed for the proposed solution might need to be changed depending upon the space available in the Versascan unit. The digital

images of the ROI used in the experiments were captured without an additional light source. Versascan is intended to be a closed box unit so an additional source of light may be needed to take digital images.

11 CONCLUSIONS AND RECOMMENDATIONS

11.1 Conclusions

The investigation showed that stereo-photogrammetry could be employed successfully to locate ROIs prior to X-ray scanning on the Lodox Statscan machine (errors were less than 13 mm). It was also shown that matching techniques such as ZNCC and ZSAD, as well as the use of circular markers, could simplify the selection of corresponding landmarks.

The developed application is simple and cost effective. It can be used for subjects of different sizes and shapes. The computer application developed can be used to calibrate cameras, capture digital images from the two cameras simultaneously and to find the position of the C-arm prior to the X-ray scanning process.

11.2 Recommendations

The Hough circle transform might not give the desired results if there are other circular objects on the object under investigation as the algorithm might detects pseudo circles in the image. In addition, the projection of a circle results in an ellipse in the image plane. Direct least square fitting of ellipses might be used to overcome this problem (Fitzgibbon *et al*., 1999). The centre of the projected circle does not necessarily correspond to the ellipse centre. A method proposed by Li et al (2004) could be used for the removal of the circle to ellipse centre projection error.

If sticking a circular marker on human subject is difficult (in case of injuries or other trauma); a hospital gown with circular markers could be designed. Such a gown might also increase the accuracy of the Hough circle transform in detecting only the retro-reflective markers and thereby avoiding the risk of selecting other circular prints or figures on the patients' clothing.

The developed computer tool could be used in applications other than X-ray. It could be used in any application that requires the 3D coordinates of visible anatomical landmarks. For example, it could be used in the detection of spinal cord disorders such as scoliosis, the deformity caused by coronal, sagittal, and axial rotation of the spine. Huysmans et al (2004) used active contours to reconstruct the 3D spinal midline for detecting scoliosis. With the proposed solution, the 3D coordinates of the landmark points indicating the spinal midline could be determined.

The shape of the human body provides information regarding health and the risk of various diseases. For example in metabolic syndrome, there is an increase in central obesity. The hip to waist ratio may be used as a measure of central obesity (Isomaa *et al*., 2001). Measuring anthropometrical distances manually is time consuming. The application could be used to measure ROIs such as waist and hip width and thickness, which along with additional information such as height, weight, and sex, could be used for predicting the risk of metabolic syndrome (Wells *et al*., 2007).

REFERENCES

Abbasi-Dezfouli, M. & Freeman, T.G. 1994, "Patch matching in stereo images based on shape", *Proceedings of SPIE*, pp. 1.

Aung, S., Ngim, R. & Lee, S. 1995, "Evaluation of the laser scanner as a surface measuring tool and its accuracy compared with direct facial anthropometric measurements", *British journal of plastic surgery*, vol. 48, no. 8, pp. 551-558.

Azad, P., Dillmann, R.,Gockel, T., 2008, Introduction to the Algorithms. In *Computer Vision Principles and Practice*. UK: Elecktor International Media BV, pp. 67-83.

Banks, J., Bennamoun, M. & Corke, P., 1997, "Non-parametric techniques for fast and robust stereo matching", *TENCON'97. IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications*, pp. 365.

Banks, J. & Corke, P. 2001, "Quantitative evaluation of matching methods and validity measures for stereo vision", *The International Journal of Robotics Research*, vol. 20, no. 7, pp. 512.

Baroni, G., Ferrigno, G., Orecchia, R. & Pedotti, A. 2000, "Real-time three-dimensional motion analysis for patient positioning verification", *Radiotherapy and Oncology*, vol. 54, no. 1, pp. 21-27.

Beets, M., P. 2007, “Distortion Correction in Lodox Statscan X-ray Images”, Masters Thesis, University of Cape Town.

Belbachir, A.N. 2009. *Smart cameras.* Springer Verlag, pp. 138.

Beningfield, S., Potgieter, J., Bautz, P., Shackleton, M., Hering, E., De Jager, G., Bowie, G., Marshall, M., Cox, G., Pagliari, G., Coetzee, N., 1999, " Evaluation of a new type of direct digital radiography machine", *South African medical Journal – Cape Town –Medical Association of South Africa*, vol. 89, pp. 1182-1187.

Beningfield, S., Potgieter, H., Nicol, A., Van As, S., Bowie, G., Hering, E. & Lätti, E. 2003, "Report on a new type of trauma full-body digital X-ray machine", *Emergency radiology*, vol. 10, no. 1, pp. 23-29.

Bennamoun, M. & Mamic, G.J. 2002. *Object recognition: fundamentals and case studies.* Springer Verlag, pp. 34-35.

Bland, J., M., Altman, D., G., 1986, "Statistical methods for assessing agreement between two methods of clinical measurement", *The Lancet*, vol. 327, no. 8476, pp. 307-310.

Blostein, S.D. & Huang, T.S. 1987, "Error analysis in stereo determination of 3-D point positions", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 752-765.

Bradski, G. & Kaehler, A., 2008. *Learning OpenCV Computer Vision with the OpenCV Library*. Beijing: Oreilly.

Braillon, C., Pradalier, C., Crowley, J.L. & Laugier, C. 2006, "Real-time moving obstacle detection using optical flow models", *Proc. Of IEEE Intelligent Vehicles Symposium*Citeseer, pp. 466.

Brown, D.C. 1971, "Close-range camera calibration", *Photogrammetric engineering*, vol. 37, no. 8, pp. 855-866.

Cao, V.T., Park, Y.Y., Shin, J.H., Lee, J.H. & Cho, H.M. 2010, "A simple method for correcting lens distortion in low-cost camera using geometric invariability",

Advanced Intelligent Computing Theories and Applications.With Aspects of Artificial Intelligence, pp. 325-333.

Cochran, S.D. & Medioni, G. 1992, "3-D surface description from binocular stereo", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 981-994.

Cootes, T.F., Taylor, C.J., Cooper, D.H. & Graham, J. 1995, "Active shape models-their training and application", *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59.

Cyganek, B. & Siebert, J.P. 2009, *An introduction to 3D computer vision techniques and algorithms*, Wiley Online Library, pp. 66-70.

De la Escalera, A. & Armingol, J.M. 2010, "Automatic Chessboard Detection for Intrinsic and Extrinsic Camera Parameter Calibration", *Sensors*, vol. 10, no. 3, pp. 2027-2044.

Dimensional Imaging Ltd
http://www.di3d.com/solutions/ [Accessed on 25 Oct 2011].

Douglas, T.S. & Mutsvangwa, T.E.M. 2010, "A review of facial image analysis for delineation of the facial phenotype associated with fetal alcohol syndrome", *American Journal of Medical Genetics Part A*, vol. 152, no. 2, pp. 528-536.

Douglas, T., S., Martinez, F., Meintjes, E., M., Vaughan, C., L., Viljoen, D., L.,2003, "Eye feature extraction for diagnosing the facial phenotype associated with fetal alcohol syndrome", *Medical and Biological Engineering and Computing* , vol. 41, no. 1, DOI: 10.1007/BF02343545.

Edwards, G.J., Cootes, T.F. & Taylor, C.J. 1998, "Face recognition using active appearance models", *Computer Vision—ECCV'98,* pp. 581.

Eriksson, G., Eklund, A.E., Torlegård, K. & Dauphin, E. 1979, "Evaluation of leg ulcer treatment with stereophotogrammetry", *British Journal of Dermatology*, vol. 101, no. 2, pp. 123-131.

Ferrario, V.F., Sforza, C., Poggio, C.E. & Schmitz, J.H. 1997, "Three-dimensional study of growth and development of the nose", *The Cleft palate-craniofacial journal*, vol. 34, no. 4, pp. 309-317.

Firefly MV CS mount camera: http://www.ptgrey.com/products/fireflymv/fireflymv.pdf [Accessed on 02 May 2011].

FDA 3626 (7/07) A guide for the submission of initial reports on diagnostic X-ray systems and their major components. Maryland:Office of compliance X-ray products branch., Food and Drug Administration, pp.51, Jan 1982 (revised). URL: http://www.fda.gov/downloads/aboutfda/reportsmanualsforms/forms/ucm081561.pdf [Accessed on 10 Nov 2010].

Fitzgibbon, A., Pilu, M. & Fisher, R.B. 1999, "Direct least square fitting of ellipses", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476-480.

Fryer, J., Chandler, J., Mitchell, H., 2007. Medicine and sport: measurement of humans. In *Application of 3D Measurement from Images*. Scotland: Whittles Publishing , pp.171-197.

Fua, P. 1991, "Combining stereo and monocular information to compute dense depth maps that preserve depth discontinuities", *Proceedings of the 12th international joint conference on Artificial intelligence*Citeseer, pp. 1292–1298.

Granshaw, S. 1980, “Bundle adjustment methods in engineering photogrammetry”, *Photogrammetric record*, vol. 10, no. 56, pp. 181-207.

Greenway, W. 2004, “Statscan Properties, and principles”. URL: http://www.aapm.org/meetings/04ss/documents/greenway1.PDF [Accessed on 12 Jan 2010].

Grobbelaar, R. & Douglas, T.S. 2007, "Stereo image matching for facial feature measurement to aid in fetal alcohol syndrome screening", *Medical engineering & physics*, vol. 29, no. 4, pp. 459-464.

Grobbelaar, R. 2005, “Stereo facial image matching to aid in Fetal alcohol syndrome screening”, Master’s Thesis, University of Cape Town.

Hannah, M.J. 1988, "Digital stereo image matching techniques", *International Archives of Photogrammetry and remote sensing*, vol. 27, no. 3, pp. 280-293.

Heikkila, J. 2000, "Geometric camera calibration using circular control points", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1066-1077.

Huysmans, T., Haex, B., Audekercke, R.,V., Sloten, J., V., Van der Perre, G., 2004, “Three-dimensional mathematical reconstruction of the spinal shape, based on active contours”, *Journal of Biomechanics*, vol. 37, no. 11, pp. 1793-1798.

Hsieh, J.W., Liao, H.Y.M., Fan, K.C., Ko, M.T. & Hung, Y.P. 1997, "Image registration using a new edge-based approach", *Computer Vision and Image Understanding*, vol. 67, no. 2, pp. 112-130.

Intille, S. & Bobick, A. 1994, "Disparity-space images and large occlusion stereo", *Computer Vision—ECCV'94*, pp. 179-186.

Irving, B., Maree, G., Hering, E. & Douglas, T. 2008, "Radiation dose from a linear slit scanning X-ray machine with full-body imaging capabilities", *Radiation Protection Dosimetry*. vol. 130, no. 4, pp. 482.

Irving, B., 2008, "Radiation dose measurement and prediction for linear slit scanning radiography", Masters Thesis, University of Cape Town.

Isomaa, B., Almgren, P., Tuomi, T., Forsén, B., Lahti, K., Nissén, M., Taskinen, M.R. & Groop, L. 2001, "Cardiovascular morbidity and mortality associated with the metabolic syndrome", *Diabetes care*, vol. 24, no. 4, pp. 683.

Kass, M., Witkin, A. & Terzopoulos, D. 1988, "Snakes: Active contour models", *International journal of computer vision*, vol. 1, no. 4, pp. 321-331.

Koeslag, A., 2002. Computer aided diagnosis of millary tb in chest x-rays. Master's thesis, University of Cape Town.

Kytö, M., Nuutinen, M. & Oittinen, P. 2011, "Method for measuring stereo camera depth accuracy based on stereoscopic vision", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 17.

Li, L., Zheng, Y., Kallergi, M. & Clark, R.A. 2001, "Improved method for automatic identification of lung regions on chest radiographs", *Academic Radiology*, vol. 8, no. 7, pp. 629-638.

Li, L.F., Feng, Z.R. & Peng, Q.K. 2004, "Detection and model analysis of circular feature for robot vision", *IEEE Machine Learning and Cybernetics, 2004*, pp. 3943.

Luhmann, T., Robson, S., Kyle, S. & Harley, I. 2006, *Close range photogrammetry: principles, techniques and applications*, Whittles Publishing Dunbeath, Scotland.

Marr, D. & Poggio, T. 1979, "A computational theory of human stereo vision", *Proceedings of the Royal Society of London.Series B.Biological Sciences,* vol. 204, no. 1156, pp. 301.

Maree, G.J., Irving, B.J. & Hering, E.R. 2007, "Paediatric dose measurement in a full-body digital radiography unit", *Pediatric radiology,* vol. 37, no. 10, pp. 990-997.

Mikhail, E.M., Bethel, J.S. & McGlone, J.C. 2001, *Introduction to modern photogrammetry*, Wiley New York, NY.

Mcvey, E.S. & Lee, J.W. 1982, "Some accuracy and resolution aspects of computer vision distance measurements", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* no. 6, pp. 646-649.

Meintjes, E., Douglas, T., Martinez, F., Vaughan, C., Adams, L., Stekhoven, A. & Viljoen, D. 2002, "A stereo-photogrammetric method to measure the facial dysmorphology of children in the diagnosis of fetal alcohol syndrome", *Medical engineering & physics,* vol. 24, no. 10, pp. 683-689.

Mitchell, H. & Newton, I. 2002, "Medical photogrammetric measurement: overview and prospects", *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 56, no. 5-6, pp. 286-294.

Miyazawa, K., Ito, K., Aoki, T., Kobayashi, K. & Nakajima, H. 2005, "An efficient iris recognition algorithm using phase-based image matching", *IEEE International Conference on Image Processing, 2005. ICIP 2005*, pp. II.

Montabone, S. 2010, *Beginning Digital Image Processing: Using Free Tools for Photographers,* Springer.

Mutsvangwa, T., Smit, J., Hoyme, H.E., Kalberg, W., Viljoen, D.L., Meintjes, E.M. & Douglas, T.S. 2009, "Design, Construction, and Testing of a Stereo-

Photogrammetric Tool for the Diagnosis of Fetal Alcohol Syndrome in Infants", *IEEE Transactions on Medical Imaging* , vol. 28, no. 9, pp. 1448-1458.

Mutsvangwa, T., 2009, "Characterization of the facial phenotype associated with fetal alcohol syndrome using stereo-photogrammetry and geometric morphometrices", PhD Thesis, University of Cape Town.

Sun Nixon, M.S. & Aguado, A.S. 2008, *Feature extraction and image processing*, Academic Press (Elsevier), Second edition.

Ohta, Y., & Kanade, T. 1983, *Stereo by intra-and inter-scanline search using dynamic programming*, Carnegie-Mellon University, Dept. of Computer Science.

Owens, R., 1997. Epipolar Geometry. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT10/node3.html [Accessed on 18 May 2010].

Photometrix 2004. *Australis users manual*. Photometrix Pty Ltd, PO Box 3023, Kew, Vic 3101, Australia.

Pitcher, R.D., van As, A.B., Sanders, V., Douglas, T.S., Wieselthaler, N., Vlok, A., Paverd, S., Kilborn, T., Rode, H. & Potgieter, H. 2008, "A pilot study evaluating the "STATSCAN" digital X-ray machine in paediatric polytrauma", *Emergency Radiology*, vol. 15, no. 1, pp. 35-42.

Plourde, F., Cheriet, F. & Dansereau, J. 2006, "Semi-automatic detection of scoliotic rib borders using chest radiographs", *Studies in health technology and informatics*, vol. 123, pp. 533-537.

Precision 3D Limited http://www.precision3d.co.uk/index.htm [Accessed on 25 Oct 2011].

Roma, N., Santos-Victor, J. & Tomée, J. 2002, A comparative analysis of cross-correlation matching algorithms using a pyramidal resolution approach .In *Empirical evaluation methods in computer vision, Edited by* Christensen, H.I. & Phillips, P.J. World Scientific Pub Co Inc. pp. 117- 125.

Savara, B.S. 1965, "Applications of photogrammetry for quantitative study of tooth and face morphology", *American Journal of Physical Anthropology,* vol. 23, no. 4, pp. 427-434.

Shapiro, L.G. & Haralick, R.M. 1985, "A metric for comparing relational descriptions", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* no. 1, pp. 90-94.

Shum, H., Ke, Q., Zhang, Z. 1999, "Efficient bundle adjustment with virtual key frames: a hierarchical approach to multi-frame structure from motion", *Computer vision and Pattern Recognition*, DOI-10.1109/CVPR.1999.784733.

Son, T.T. & Mita, S., 2004, "A new approach using the viterbi algorithm in stereo correspondence problem", *IEEE International Conference on Systems, Man and Cybernetics,* pp. 3016.

Sun, C. & Peleg, S. 2004, "Fast panoramic stereo matching using cylindrical maximum surfaces", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,*vol. 34, no. 1, pp. 760-765.

Takamoto, T. & Schwartz, B. 1986, "Three-dimensional mapping of retinal vessels in ophthalmology", *International Archives of Photogrammetry and Remote Sensing,* vol. 26, no. 5, pp. 349–354.

Tamron CCTV lens specifications: http://www.tamron.co.jp/en/data/cctv/13vm2812as2.html [Accessed on 02 May 2011].

Tredwell, S., Booth, K., Reilly, C., Sawatzky, B. & Jang, S. 1995, Three-Dimensional analysis of scoliosis surgery using stereophotogrammetry. In D'Amico, M., Merolli, A., Santambrogio, G., C., (eds). *Three Dimensional Analysis of Spinal Deformities,* Amsterdam, The Netherlands: IOS, pp. 92-100.

Triggs, B., McLauchlan, P., F., Hartely, R., I., Fitzgibbon, A., W. 1999 "Bundle adjustment – A modern Synthesis", *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*, pp. 298-372.

Utkarsh. Two major physical defects in cameras. URL: http://www.aishack.in/2010/07/two-major-physical-defects-in-cameras/ [Accessed on 02 May 2011].

Van Ginneken, B., Stegmann, M.B. & Loog, M. 2006, "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database", *Medical image analysis,* vol. 10, no. 1, pp. 19-40.

Vaughan, C.L. 2008, "Digital X-rays come of age", *South African Medical Journal,* vol. 96, no. 7, pp. 610.

Weinberg, S.M., Naidoo, S., Govier, D.P., Martin, R.A., Kane, A.A. & Marazita, M.L. 2006, "Anthropometric precision and accuracy of digital three-dimensional photogrammetry: Comparing the Genex and 3dMD imaging systems with one another and with direct anthropometry", *Journal of Craniofacial Surgery,* vol. 17, no. 3, pp. 477.

Wells, J., CK., Treleavan, P., Cole, T., J.2007, "BMI compared with 3-dimensional body shape: the UK National Sizing Survey", *American Journal of Clinical Nutrition*, vol. 85, no. 2, pp. 419-425.

Yu, C., 2010, “Canny Edge Detection and Hough Transform”. (ppt) http://www.indiana.edu/~dll/B657/B657_lec_hough.pdf [Accessed on 10 Feb 2011].

Yue, Z., Goshtasby, A. & Ackerman, L.V. 1995, "Automatic detection of rib borders in chest radiographs", *IEEE Transactions on Medical Imaging*, vol. 14, no. 3, pp. 525-536.

Zabih, R. & Woodfill, J. 1994, "Non-parametric local transforms for computing visual correspondence", *Computer Vision—ECCV'94*, pp. 151-158.

Zhang, Z. 2004, "Camera calibration with one-dimensional objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 892-899.

Zhang, Z. 2000, "A flexible new technique for camera calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334.

3dMD Inc
http://www.3dmd.com/ [Accessed on 25 Oct 2011].